Auswirkungen der Haftungsregelungen auf die Prüfungsqualität: Eine kritische Würdigung

Bachelorarbeit

Max Mustermann

Musterstraße 12, 12345 Musterstadt

E-Mail: max.mustermann@example.com

Matrikelnummer: 1234567

Studiengang, 3. Fachsemester

Universität Musterstadt

Dozent: Prof. Dr. Anna Beispiel

Titel der Veranstaltung: Beispiel

Semester: Wintersemester 2025/26

Inhaltsverzeichnis

Abkürzungsverzeichnis IV

Abstract V

1. Einleitung 1

1.1 Problemstellung und Relevanz 1

1.2 Zielsetzung, wissenschaftlicher Beitrag und Abgrenzung 1

1.3 Forschungsfrage(n) und Hypothesen 2

1.4 Methodik und Aufbau der Arbeit 3

2. Grundlagen der Abschlussprüfung 3

2.1 Prinzipal–Agent-Theorie und Informationsasymmetrien 3

2.2 Ziele und Zwecke der Abschlussprüfung 4

2.3 Prüfungsrisiken und Anreizprobleme 6

3. Grundlagen der Prüfungsqualität 7

3.1 Begriffsdefinition(en) und Dimensionen von Prüfungsqualität 7

3.2 Messung von Prüfungsqualität 8

3.3 Rechtlicher/berufsständischer Qualitätsrahmen 10

4. Zivilrechtliche Haftung und Regulierungsmechanismen 11

4.1 Haftungsarchitektur in Deutschland 11

4.2 Wirkungsmechanismen: Haftung - Prüfungsanstrengung, Honorar, Mandatsannahme, Dokumentation 12

4.3 Ländervergleich/Institutionenvergleich als Kontrastfolie 14

5. Systematischer Literaturüberblick 15

5.1 Such- und Auswahlstrategie 15

5.2 Kodierschema/Literaturmatrix: Variablen, Qualitätsmaße, Haftungsmaß, Kontext 16

5.3 Evidenzblock A: Befunde „Haftung erhöht Qualität“ 17

5.4 Evidenzblock B: Befunde „Haftung erhöht Qualität nicht / Nebenwirkungen dominieren“ 18

5.5 Evidenzblock C: Interdependenzen und gemischte Effekte 20

6. Kritische Würdigung 21

6.1 Kritische Diskussion der Literatur 21

6.2 Eigene Bewertung und Implikationen 22

7. Fazit und Ausblick 23

7.1 Beantwortung der Forschungsfrage(n) und Rückbezug auf Hypothesen 23

7.2 Limitationen der Arbeit 24

7.3 Ausblick 25
Literaturverzeichnis 27
Anhang A: Kodierschema 31
Anhang B: Literaturmatrix 33

Abkürzungsverzeichnis

BGB	Bürgerliches Gesetzbuch
Big 4	Vier größte Wirtschaftsprüfungsgesellschaften
FISG	Finanzmarktintegritätsstärkungsgesetz
GCM	Going-Concern-Modifikation(en) (im Bestätigungsvermerk)
HGB	Handelsgesetzbuch
IESBA	International Ethics Standards Board for Accountants
ISA	International Standards on Auditing
NAS	Non-Audit Services (Nichtprüfungsleistungen)
PIE	Public-Interest Entity (Unternehmen von öffentlichem Interesse)
QM	Qualitätsmanagement
QS	Qualitätssicherung

Abstract

Diese Bachelorarbeit untersucht, wie zivilrechtliche Haftungsregelungen die Qualität der Abschlussprüfung beeinflussen. Ausgangspunkt sind Bilanzskandale und der daraus folgende politische Druck, Haftung und Unabhängigkeit als Qualitätshebel zu verschärfen, obwohl zugleich Risiken für Versicherbarkeit, Marktstruktur und Nachwuchs diskutiert werden. Die Arbeit ist als systematische Literaturrecherche angelegt. Sie verbindet ökonomische Anreizmechanismen (Prinzipal–Agent, Quasi-Renten, Litigation- und Reputationsrisiko) mit einem mehrdimensionalen Qualitätsverständnis (Output, Prozess, Wahrnehmung) und ordnet empirische Proxies kritisch ein. Im Vergleich verschiedener Haftungsregime zeigt die Literatur kein eindeutiges „mehr Haftung = bessere Qualität“. Häufig reagieren Prüfer eher über höhere Honorare oder risikoselektive Mandatsannahme als über mehr Prüfungsanstrengung; Reputation und Enforcement können Haftungsgrenzen teilweise ersetzen. Insgesamt sprechen die Befunde für kalibrierte, risikoorientierte Haftungsmodelle und für Evaluierungen, die Qualität, Konzentration und Versicherbarkeit gemeinsam messen. Die Synthese umfasst deutschen Haftungsrahmen mit Haftungsobergrenzen sowie Lösungen in Frankreich und im Vereinigten Königreich. Methodische Grenzen ergeben sich aus Proxy-Sensitivität, Endogenität und begrenzter Evidenz zu Rotation, Joint Audits und Regulierungsbündeln.

1. Einleitung

1.1 Problemstellung und Relevanz

Bilanzskandale stoßen regelmäßig Reformen der Abschlussprüfung an. Nach dem Wirecard-Fall gerieten in Deutschland auch Wirtschaftsprüfer stärker unter Regulierungsdruck; Pfaff und Ruffing-Straube verweisen dabei auf Parallelen zur Enron-Krise und den anschließenden US-Reformen (Pfaff & Ruffing-Straube, 2021, S. 239). Im Finanzmarktintegritätsstärkungsgesetz wird diese Dynamik sichtbar: Der Gesetzgeber verbindet Unabhängigkeitsanforderungen und Haftungsfragen ausdrücklich mit dem Ziel, die Prüfungsqualität zu verbessern (Institut der Wirtschaftsprüfer in Deutschland e. V. [IDW], 2021, S. 5).

Ökonomisch ist der Ansatz plausibel, weil Haftungs- und Reputationsrisiken das Verhalten von Prüfern beeinflussen. Die Prüfungsanstrengung hängt nicht nur vom Berufsethos, sondern auch vom Litigation Risk und vom Reputationsrisiko ab (Pfaff & Ruffing-Straube, 2021, S. 240). Zugleich ist strittig, ob eine Haftungsverschärfung tatsächlich zu höherer Qualität führt. Das IDW bezweifelt einen Qualitätseffekt und betont Grenzen der Versicherbarkeit, insbesondere bei faktisch unbegrenzter Haftung bereits bei grober Fahrlässigkeit (IDW, 2021, S. 1, 5). Dann kann der Qualitätsanreiz in defensive Risikovermeidung umschlagen.

Hinzu treten Marktfolgen. Höhere Haftungsrisiken können kleinere und mittelständische Praxen verdrängen und die Konzentration verstärken (IDW, 2021, S. 5). Das ist besonders relevant, weil der Prüfungsmarkt bereits stark konzentriert ist; für Deutschland wird ein sehr hoher Marktanteil der größten Anbieter berichtet (Maca & Tirdea, 2025, S. 247). Zusammen mit weiteren Instrumenten der EU-Auditreform wie Rotation und Beschränkungen von Nichtprüfungsleistungen (Maca & Tirdea, 2025, S. 246) stellt sich damit auch eine Versorgungsfrage: ob Unternehmen verlässlich Prüfer finden und der Berufsstand ausreichend Nachwuchs gewinnt (IDW, 2021, S. 5, 16).

1.2 Zielsetzung, wissenschaftlicher Beitrag und Abgrenzung

Vor diesem Hintergrund untersucht die Arbeit, ob zivilrechtliche Haftungsregeln als Hebel zur Steigerung der Prüfungsqualität wirken oder ob Nebenwirkungen die intendierten Effekte neutralisieren. Zweifel werden sowohl in berufsständischen Stellungnahmen als auch in reformkritischen Beiträgen formuliert (IDW, 2021, S. 14; Deutsches Aktieninstitut, 2021, S. 6).

Ziel ist eine systematische Literaturarbeit, die Befunde zur Wirkung von Haftung auf Prüfungsqualität vergleichbar macht. Im Mittelpunkt stehen Mechanismen über Prüfungsanstrengung und Skepsis, über Honorare und Risikoprämien sowie über Selektionseffekte wie Mandatsvermeidung und Marktkonzentration (Pfaff & Ruffing-Straube, 2021, S. 240–241; IDW, 2021, S. 5). Der Beitrag liegt in einer strukturierten Synthese gegensätzlicher Positionen, einschließlich Reformargumenten für höhere Haftungsobergrenzen (Langenbucher et al., 2020, S. 17) und Hinweisen auf reputationsbasierte Disziplinierung trotz begrenzter Haftung (Downar et al., 2021, S. 3). Abgegrenzt ist die Arbeit durch den Verzicht auf eigene Empirie; sie ordnet ausschließlich die bereitgestellten Studien sowie Rechts- und Policy-Quellen in die langfristig stabile, aber umstrittene Haftungsarchitektur ein (Kuhmann, 2025, S. 21).

1.3 Forschungsfrage(n) und Hypothesen

Die zentrale Forschungsfrage lautet: Wie wirken zivilrechtliche Haftungsregelungen für Abschlussprüfer auf die Prüfungsqualität und die dafür verwendeten Qualitätsproxies? Der Ausgangspunkt ist kein Automatismus „mehr Haftung = bessere Prüfung“, weil die Literatur auch gegenläufige Kanäle beschreibt (Pfaff & Ruffing-Straube, 2021, S. 240–241).

Drei Unterfragen strukturieren die Analyse: (1) Führt höheres Haftungs- bzw. Klagerisiko zu mehr Prüfungsanstrengung oder vor allem zu höheren Honoraren (Pfaff & Ruffing-Straube, 2021, S. 240–241)? (2) In welchem Ausmaß schwächen Haftungsobergrenzen die Anreize, und kann Reputation diesen Effekt teilweise ersetzen (Downar et al., 2021, S. 3; Langenbucher et al., 2020, S. 14)? (3) Welche Nebenwirkungen treten über Marktaustritte, Konzentration und defensive Haftungsvermeidung auf, auch im Zusammenspiel mit weiteren Reforminstrumenten (IDW, 2021, S. 5; Sethe, 2012, S. 128; Ruhwedel et al., 2021, S. 16)?

Daraus folgen konzeptionelle Erwartungen: H1 unterstellt einen positiven Qualitätseffekt über mehr Effort unter höherem Haftungsdruck (Völker, 2021, S. 122; Pfaff & Ruffing-Straube, 2021, S. 240). H2 erwartet eine stärkere Preis- als Effort-Reaktion (Pfaff & Ruffing-Straube, 2021, S. 241). H3 nimmt an, dass Haftungscaps Anreize dämpfen, bei hohen Reputationskosten jedoch weniger stark (Downar et al., 2021, S. 3). H4 erwartet, dass starke Haftungsverschärfungen über Marktstruktur und defensive Strategien Qualitätsgewinne neutralisieren oder umkehren können (IDW, 2021, S. 5; Sethe, 2012, S. 128).

1.4 Methodik und Aufbau der Arbeit

Die Arbeit ist als strukturierte Literaturrecherche konzipiert und kombiniert wissenschaftliche Fachliteratur (Artikel, Monografien, Dissertationen) mit Policy- und Rechtsquellen, weil sich die Debatte an der Schnittstelle von ökonomischen Anreizen und rechtlicher Durchsetzung bewegt (IDW, 2021, S. 5; Deutscher Bundestag, Finanzausschuss, 2021, S. 16). Berücksichtigt werden Quellen, die Haftung oder eng verwandte Regulierungsinstrumente (z. B. Rotation, NAS-Beschränkungen) explizit mit Prüfungsqualität oder Qualitätsproxies verknüpfen (Maca & Tirdea, 2025, S. 246). Vergleichskontexte werden einbezogen, wenn sich institutionelle Bedingungen wie Aufsichtsdichte unterscheiden (Martani et al., 2021, S. 5) oder Reputation als Sanktionsmechanismus bei begrenzter Haftung wirkt (Downar et al., 2021, S. 3).

Alle Quellen werden in einer Literaturmatrix kodiert: Haftungs-/Regulierungsvariable, Qualitätsmaß, Argumenttyp/Identifikationslogik, vermuteter Wirkmechanismus sowie Kontextfaktoren (Enforcement, Marktstruktur, Versicherbarkeit). Methodisch wichtig ist dabei, dass Honorare zugleich Aufwand und Risikoanteile abbilden können und daher als Qualitätsproxy nur begrenzt eindeutig sind (Choi et al., 2010, S. 2; Joha, 2018, S. 148). Zusätzlich wird erfasst, dass Studien je nach Qualitätsmaß (z. B. accruals) zu unterschiedlichen Ergebnissen kommen können (García-Blandón et al., 2021, S. 1393). Der Aufbau folgt der Kette Haftungsregime–Anreize–Marktreaktionen–Qualität und führt über Theorie (Kapitel 2–3) und Haftungsarchitektur (Kapitel 4) zur Literatursynthese (Kapitel 5), der kritischen Würdigung (Kapitel 6) und den Schlussfolgerungen (Kapitel 7), wobei Interdependenzen im Reformbündel als Querschnitt sichtbar bleiben (Ruhwedel et al., 2021, S. 6).

2. Grundlagen der Abschlussprüfung

2.1 Prinzipal–Agent-Theorie und Informationsasymmetrien

Zwischen den Adressaten eines Jahresabschlusses und der Unternehmensleitung besteht ein strukturelles Informationsgefälle. Die Leitung erstellt und steuert die Rechnungslegung, während externe Adressaten auf veröffentlichte Informationen angewiesen sind. Daraus ergibt sich die Gefahr, dass Fehler oder Regelverstöße nicht erkannt werden oder erst spät sichtbar werden.

Die gesetzliche Abschlussprüfung ist vor diesem Hintergrund als Mindestmechanismus zur Begrenzung dieser Asymmetrie angelegt. Der Abschlussprüfer hat die Prüfung so zu planen, dass „Unrichtigkeiten und Verstöße (…) … bei gewissenhafter Berufsausübung erkannt werden“ (§ 317 Abs. 1 Satz 3 HGB; Behringer, 2019, S. 137). Diese Vorgabe beschreibt keinen

Garantieschutz, sondern einen Sorgfaltsmaßstab, an dem sich die erwartete Leistung des Prüfers orientiert.

Gerade weil die Prüfung keine vollständige Sicherheit liefern kann, bleibt die Erwartung der Öffentlichkeit jedoch häufig weitergehend. Behringer (2019, S. 138) betont, dass vollständige Sicherheit nicht erreichbar ist und damit eine Erwartungslücke fortbesteht. In der Praxis bedeutet das: Selbst eine ordnungsgemäß durchgeführte Prüfung kann von außen als unzureichend wahrgenommen werden, wenn später ein Skandal bekannt wird.

Die verbleibende Unsicherheit zwingt den Prüfer zu einer Haltung, die Fehler auch gegen Widerstände aufdecken soll. Behringer (2019, S. 138) verweist darauf, dass ISA 210 eine skeptische Grundhaltung während der Prüfung fordert. Diese Anforderung lässt sich als Versuch verstehen, Informationsasymmetrien nicht nur formal, sondern auch verhaltensbezogen zu adressieren und die Erwartungslücke zu verkleinern.

Diese Konstellation macht verständlich, warum Haftung und Reputation als disziplinierende Gegenkräfte diskutiert werden. Pfaff und Ruffing-Straube (2021, S. 240) verorten die Prüfungsanstrengung unter anderem in der Erwartung, verklagt zu werden, sowie im Reputationsrisiko. Wird ein Abschlussfehler öffentlich, verschiebt sich für den Prüfer nicht nur das rechtliche Risiko, sondern auch das Risiko künftiger Mandatsverluste.

Für die Analyse von Unabhängigkeit und Qualität reicht es daher nicht, allein auf formale Prüfregeln zu schauen. Joha (2018, S. 19) verweist darauf, dass Modelle zur Unabhängigkeit zusätzliche Aspekte wie Haftung, Reputation und den Informationswert der Jahresabschlussinformation für Adressaten berücksichtigen sollten. Damit lässt sich festhalten, dass Informationsasymmetrien und Erwartungslücken den Boden für haftungs- und reputationsgetriebene Anreize bilden und die spätere Bewertung von Haftungsverschärfungen nur über diese Wirkkanäle konsistent möglich ist.

2.2 Ziele und Zwecke der Abschlussprüfung

Die Abschlussprüfung lässt sich in dieser Arbeit als Antwort auf Informationsasymmetrien verstehen. Ihre Ziele erklären, warum Haftungs- und Unabhängigkeitsregeln überhaupt als Qualitätshebel diskutiert werden.

Im Kern steht der gesetzliche Assurance-Anspruch. Nach § 317 Abs. 1 Satz 3 HGB ist die Prüfung so zu planen, dass „Unrichtigkeiten und Verstöße (…) … bei gewissenhafter Berufsausübung erkannt werden“ (Behringer, 2019, S. 137). Damit wird kein

Erfolgsversprechen abgegeben, sondern ein Sorgfaltsmaßstab, an dem Verhalten und Dokumentation des Prüfers gemessen werden.

Zu dieser Assurance gehört eine bestimmte Arbeitshaltung. Behringer betont, dass ISA 210 eine skeptische Grundhaltung während der Prüfung verlangt. Skepsis ist damit nicht nur Technik, sondern soll als Prozessmerkmal helfen, Fehlerhinweise zu suchen und zu bewerten.Gleichzeitig zielt Abschlussprüfung auf Vertrauen, stößt aber an Grenzen. Vollständige Sicherheit ist nicht erreichbar, weshalb eine Erwartungslücke der Öffentlichkeit bestehen bleibt (Behringer, 2019, S. 138). Gerade diese Lücke erklärt, warum nach Skandalen häufig nach stärkeren Sanktionen oder Haftungsregeln gerufen wird.

Neben Assurance und Vertrauen hat die Prüfung eine Versicherungs- und Signalfunktion. Joha beschreibt, dass Nachfrage nach Prüfung auch als Versicherungsbedürfnis verstanden werden kann und verweist auf die Deep-Pockets-Theorie (Joha, 2018, S. 40). Adressaten erwarten dann nicht nur bessere Information, sondern auch einen potenziell haftungsfähigen „Zahler“ im Schadensfall.

Diese Logik wirkt auch über die Wahl der Prüfungsgesellschaft. Für Börsengänge wird die Entscheidung für eine große Prüfungsgesellschaft laut Joha häufig als Versicherungssignal interpretiert (Joha, 2018, S. 29). Die Prüferwahl sendet damit ein extern beobachtbares Zeichen, das an Reputation und vermutete Haftungskapazität anknüpft.

Die Versicherungs- und Vertrauenskomponente erhöht zugleich die Sensibilität für Prüfungsfehler. Kommt es zu einem Fehlurteil, drohen Klagen und Reputationsverluste; in der Literatur wird deshalb betont, dass Prüfungsanstrengung auch vom Litigation Risk und vom Reputationsrisiko abhängt (Pfaff & Ruffing-Straube, 2021, S. 240). Der Zweck der Prüfung ist damit nicht nur Informationsprüfung, sondern auch ein Mechanismus, der über mögliche Sanktionen Verhalten beeinflusst und die Kosten der Prüfung mitprägen kann. Die verbleibende Erwartungslücke kann diesen Druck erhöhen und die Prüfung für viele Adressaten stärker als Absicherung erscheinen lassen (Behringer, 2019, S. 138; Joha, 2018, S. 40).

Die drei Zwecke stehen nicht isoliert nebeneinander, sondern beeinflussen die Anreize im Prüfungsmarkt. Wenn Prüfung zugleich Sorgfalt nachweisen, Vertrauen stützen und als Versicherung dienen soll, rücken Haftungs- und Reputationsrisiken als Steuerungsmechanismen in den Vordergrund. Vor diesem Hintergrund wird deutlich, dass spätere Analysen zur Haftung immer prüfen müssen, ob Regulierung primär die Assurance stärkt oder vor allem die Versicherungs- und Signalrolle verändert.

2.3 Prüfungsrisiken und Anreizprobleme

Die Abschlussprüfung soll Vertrauen schaffen, bleibt aber wegen begrenzter Prüfungssicherheit konfliktanfällig. In diesem Umfeld wird Prüfungsqualität nicht nur durch Fachkompetenz, sondern auch durch Anreize bestimmt, die das Verhalten des Prüfers im Mandat steuern.

Ein zentrales Risiko liegt in der Unabhängigkeit. Behringer (2019, S. 139) betont, dass professionelle Skepsis besonders durch wirtschaftliche Abhängigkeiten eingeschränkt werden kann. Damit wird ein Spannungsfeld sichtbar: Der Prüfer soll kritisch bleiben, ist aber zugleich auf den Mandanten als Einnahmequelle angewiesen.

Ökonomisch lässt sich diese Abhängigkeit über mandantenspezifische Quasi-Renten fassen. Choi et al. (2010, S. 7–8) argumentieren, dass solche Quasi-Renten einen Anreiz schaffen, die Unabhängigkeit zu kompromittieren; ohne diese Renten wäre der Prüfer gegenüber dem jeweiligen Mandanten vollständig unabhängig. Quasi-Renten sind damit weniger ein Randphänomen als ein Mechanismus, der das Mandatsverhältnis strukturell prägt.

Eng damit verbunden ist der Honorarkanal. Nach Choi et al. (2010, S. 2) spiegeln Prüfungshonorare zugleich Aufwand, Haftungsrisiko und einen Normalgewinn wider. Hohe Honorare können daher zweierlei bedeuten: mehr eingesetzte Ressourcen oder eine stärkere Risikoprämie.

Gleichzeitig warnen Choi et al. (2010, S. 2-3), dass überhöhte Honorare Anreize setzen können, Mandantendruck nachzugeben. Empirisch sind abnormal hohe Honorare negativ mit Prüfungsqualität verbunden.

Haftungs- und Reputationsrisiken wirken als Gegenkräfte zu dieser ökonomischen Bindung. Pfaff und Ruffing-Straube (2021, S. 240) beschreiben, dass die Prüfungsanstrengung unter anderem vom Klagerisiko und vom Reputationsrisiko beeinflusst wird. Der Prüfer kann auf steigende Risiken jedoch auch primär über Honoraraufschläge reagieren, statt den Aufwand zu erhöhen, was die Zuordnung von Ursache und Wirkung erschwert (Pfaff & Ruffing-Straube, 2021, S. 241).

Der Honorar- und Quasi-Rentenmechanismus ist auch für die Interpretation von Qualitätsmessung wichtig. Wenn Gebühren sowohl Effort als auch Risiko abbilden, können steigende Honorare nicht automatisch als Signal für bessere Prüfung gelten. Gleichzeitig können starke ökonomische Bindungen die skeptische Grundhaltung unterlaufen, die für eine gewissenhafte Prüfung gefordert ist (Behringer, 2019, S. 139), und muss in Studien vorsichtig gelesen werden.

Für die Analyse ist deshalb zu trennen, ob Honorare als Qualitätsproxy gelesen werden oder als Preis für Risiko. Joha (2018, S. 30) weist darauf hin, dass höhere Honorare auch durch eingepreiste Risikoprämien aus Haftungs- und Reputationsrisiken getrieben sein können. In Kombination mit der Forderung, Unabhängigkeit unter Einbezug von Haftung und Reputation zu denken (Joha, 2018, S. 19), wird deutlich, dass Haftungsregeln spätere Qualitätseffekte nur über dieses Anreizgefüge entfalten können.

3. Grundlagen der Prüfungsqualität

3.1 Begriffsdefinition(en) und Dimensionen von Prüfungsqualität

Prüfungsqualität ist das Bindeglied zwischen dem gesetzlichen Prüfungsauftrag und dem Vertrauen der Adressaten in den Jahresabschluss. Für die spätere Analyse von Haftungsanreizen braucht es daher eine Definition, die sowohl Ergebnisse als auch Entstehungsbedingungen und Beurteilungsperspektiven trennt. Sie bildet die Basis der Messdiskussion.

Aus Output-Sicht knüpft die klassische Definition an die Wahrscheinlichkeit an, dass ein Prüfer wesentliche Fehler entdeckt und anschließend auch offenlegt. DeAngelo fasst dies als „Wahrscheinlichkeit … dass Fehler … erkannt … und auch berichtet werden“ zusammen (Brähler et al., 2011, S. 4). Output-Qualität liegt damit nicht erst in der Existenz von Prüfungsnachweisen, sondern in einem Prüfungsergebnis, das sich in der Berichterstattung des Prüfers niederschlägt.

Diese Ergebnisdimension lässt sich zugleich normativ an den gesetzlichen Mindeststandard anbinden. Der Abschlussprüfer hat die Prüfung so zu planen, dass „Unrichtigkeiten und Verstöße (…) … bei gewissenhafter Berufsausübung erkannt werden“ (Behringer, 2019, S. 137). Prüfungsqualität bedeutet in diesem Sinne, dass der Output nicht zufällig zustande kommt, sondern an einem Sorgfaltsmaßstab gemessen wird, der die Erkennungsleistung als Pflicht setzt.

Die Output-Definition umfasst zwei Schritte: Entdecken und Berichten. Das ist relevant, weil eine Entdeckung ohne klare Berichterstattung ihren Nutzen für Adressaten verfehlt. Hier zeigt sich die Nähe zur Unabhängigkeit, die durch wirtschaftliche Abhängigkeiten belastet sein kann (Behringer, 2019, S. 139).

Prozessqualität beschreibt, wie dieser Output zustande kommt. Sie umfasst insbesondere die geforderte Haltung und Arbeitsweise im Prüfungsablauf. Ein zentraler Anker ist die professionelle Skepsis, die ISA 210 während der Abschlussprüfung verlangt (Behringer, 2019,

S. 138). Skepsis meint hier nicht Misstrauen als Selbstzweck, sondern eine konsequente Prüfungsperspektive, die auf das Auffinden von Unrichtigkeiten ausgerichtet bleibt.

Gleichzeitig ist Prozessqualität anfällig für Anreizprobleme. Behringer (2019, S. 139) betont, dass wirtschaftliche Abhängigkeiten die professionelle Skepsis besonders stark einschränken können. Damit wird Prozessqualität zur Schnittstelle zu den in Kapitel 2 diskutierten Unabhängigkeitsrisiken: Auch bei unverändertem gesetzlichen Auftrag kann die tatsächliche Prüfungsdurchführung von ökonomischen Bindungen beeinflusst werden.

Wahrnehmungsqualität schließlich betrifft, wie Prüfungsqualität von relevanten Akteuren eingeschätzt wird und wie stark diese Einschätzung vom Umfeld geprägt ist. Völker (2021, S. 108) beschreibt, dass in Regionen mit starken Rechtsinstitutionen und entwickelten Märkten Haftungs- und Reputationsrisiken höher sind und Prüfungspartner deshalb die Prüfungsqualität grundsätzlich höher ansetzen. Wahrnehmung hängt damit nicht nur vom einzelnen Mandat ab, sondern auch von institutionellen Rahmenbedingungen, die Erwartungen und Verhaltensstandards prägen.

Auch Corporate-Governance-Gremien sind Teil dieser Beurteilungsdimension. Der Prüfungsausschuss soll sich ein Bild über die Sorgfalt der Arbeit des Abschlussprüfers machen (Hönsch & Kaspar, 2021, S. 80). Diese Einschätzung wirkt rückkoppelnd, weil sie Auswahlentscheidungen, Kommunikation und gegebenenfalls Nachfragen an den Prüfer beeinflusst.

Insgesamt zeigt die Dreiteilung, dass Prüfungsqualität weder allein über den Prüfungsbericht noch allein über Prüfungsprozesse erfasst werden kann, sondern stets das Zusammenspiel aus Output, Prozess und Wahrnehmung im jeweiligen institutionellen Kontext beschreibt als späterer Maßstab für Haftungs- und Regulierungswirkungen.

3.2 Messung von Prüfungsqualität

Prüfungsqualität ist in Kapitel 3.1 als Output-, Prozess- und Wahrnehmungsqualität gefasst. Empirische Studien müssen diese abstrakten Dimensionen in beobachtbare Größen übersetzen. Dabei entsteht das Grundproblem, dass viele Indikatoren zugleich Verhalten des Prüfers und Umfeldrisiken abbilden und deshalb mehrdeutig sind. Ein Indikator muss deshalb immer als Annäherung verstanden werden, nicht als direkte Beobachtung der in 3.1 beschriebenen Qualität in allen drei Dimensionen.

Viele Messansätze konzentrieren sich auf die Output-Dimension, weil Prozessqualität (z. B. Skepsis, Umfang der Prüfungshandlungen) meist nicht öffentlich beobachtbar ist. Gerade im

Kontext von Haftung ist diese Einschränkung relevant: Wenn Haftungsdruck vor allem Dokumentation und interne Kontrollen verstärkt, kann das in externen Output-Proxies nur verzögert oder gar nicht sichtbar werden. Für die Literaturauswertung bedeutet das, dass Nullbefunde nicht automatisch fehlende Verhaltensänderungen anzeigen, sondern auch Messgrenzen widerspiegeln können.

Ein häufig genutzter Proxy sind Prüfungshonorare. Honorare spiegeln nicht nur Aufwand, sondern auch Haftungs- bzw. Prozessrisiken und einen Normalgewinn wider (Choi et al., 2010, S. 2). Steigende Fees können daher auf mehr Prüfungsarbeit hindeuten, aber ebenso auf eine eingepreiste Risikoprämie, etwa für mögliche Haftungs- und Reputationsschäden (Völker, 2021, S. 20; Joha, 2018, S. 148). Zudem zeigen Befunde, dass abnormal hohe Honorare mit geringerer Prüfungsqualität zusammenhängen und überhöhte Fees Anreize schaffen können, Mandantendruck nachzugeben (Choi et al., 2010, S. 2–3). Damit sind Honorare als Qualitätsmaß besonders anfällig für Endogenität, weil das Honorar selbst Teil des Anreizsystems ist.

Die Aussagekraft von Fees ist außerdem kontextabhängig. Für Deutschland wird argumentiert, dass das Haftungsrisiko vergleichsweise gering sei und Honorare daher weniger stark durch Risikoprämien verzerrt würden, was tendenziell eher Rückschlüsse auf Qualitätsunterschiede zulässt (Völker, 2021, S. 272). Diese Perspektive bleibt jedoch eine Einschränkung: Auch bei niedrigem Litigation Risk können Honorare mehrere Komponenten bündeln und damit Interpretationsspielräume eröffnen (Joha, 2018, S. 148).

Als outputnaher Indikator werden Going-Concern-Modifikationen diskutiert, weil sie ein direktes Ergebnis der Prüfung darstellen (Günther, 2020, S. 1). Gleichzeitig kann zwischen GCM und Klagerisiko eine wechselseitige Beziehung bestehen. Wird diese Endogenität berücksichtigt, sinkt die Klagewahrscheinlichkeit nach einer GCM signifikant, sodass einfache Zusammenhänge leicht fehlinterpretiert werden (Günther, 2020, S. 146).

Haftungsfälle aufgrund fehlerhafter Prüfungen sind ein weiteres Surrogat. Ihr Vorteil liegt in hoher Messgenauigkeit, weil ein Haftungsfall die Verbindung zwischen Prüfungsfehler und Schaden konkretisiert (Günther, 2020, S. 53). Empirisch sind solche Ereignisse jedoch selten, was die statistische Auswertbarkeit und die Generalisierbarkeit einschränkt (Günther, 2020, S. 53).

Schließlich dominiert in der Forschung häufig die Messung über Accruals; zugleich zeigen unterschiedliche Ergebnisse zwischen discretionary accruals und earnings benchmarks eine zentrale Lücke und eine starke Proxy-Sensitivität (García-Blandón et al., 2021, S. 1393). Vor diesem Hintergrund wird deutlich, dass jede spätere Bewertung von Haftungswirkungen nur

dann tragfähig ist, wenn Befunde stets zusammen mit dem verwendeten Qualitätsproxy, dessen Mehrdeutigkeit und möglichen Endogenitäten gelesen werden.

3.3 Rechtlicher/berufsständischer Qualitätsrahmen

Der rechtliche und berufsständische Qualitätsrahmen konkretisiert, was im Sinne dieser Arbeit als „sorgfältige“ Prüfung gelten soll. Er ist damit der Maßstab, an dem spätere Aussagen über Haftungsanreize und tatsächliche Prüfungsqualität gespiegelt werden.

Im HGB wird Prüfungsqualität zunächst über Pflichten definiert. Abschlussprüfer sind zu einer gewissenhaften und unparteiischen Durchführung verpflichtet; bei vorsätzlicher oder fahrlässiger Pflichtverletzung entsteht eine Ersatzpflicht (Deutscher Bundestag, Wissenschaftliche Dienste, 2020, S. 5). Diese Kopplung von Sorgfaltsstandard und Haftungsfolge macht deutlich, dass Qualität nicht nur eine technische, sondern auch eine rechtlich sanktionierte Erwartung ist.

Eng damit verbunden ist die Unabhängigkeit. Für die gesetzliche Abschlussprüfung wird Unabhängigkeit als Grundanforderung genannt (Hönsch & Kaspar, 2021, S. 76). In den vorangegangenen Abschnitten wurde wirtschaftliche Abhängigkeit als typische Gefährdung professioneller Skepsis eingeordnet; der normative Rahmen setzt hier die Grenze, ab der ein Mandat oder einzelne Leistungen nicht mehr vertretbar sind.

Neben individuellen Pflichten verlangt der Rahmen organisatorische Qualitätssicherung. Das IDW beschreibt, dass Praxen zusätzliche Qualitätsziele festlegen können, um Haftungsrisiken zu begrenzen (IDW, 2026, S. 10). Schon vor Annahme oder Fortführung eines Mandats soll eine Integritäts- und Risikoanalyse erfolgen, ausdrücklich auch mit Blick auf Haftungs- und Reputationsrisiken (IDW, 2026, S. 22). Werden Unabhängigkeitsrisiken erkannt, kann dies bis zur Ablehnung oder Kündigung des Auftrags reichen (IDW, 2026, S. 60).

Ein zentraler QS-Baustein ist die Dokumentation. Berufsangehörige sollen ihre Arbeit so festhalten, dass sie nachvollziehbar und nachweisbar bleibt (International Ethics Standards Board for Accountants [IESBA], 2018, S. 26). Ergänzend betont das IDW die Aufbewahrung von Prüfungsunterlagen grundsätzlich über zehn Jahre ab Datum des Bestätigungsvermerks (IDW, 2026, S. 54), was den haftungsbezogenen Charakter der Nachweisführung unterstreicht.

Schutzmaßnahmen umfassen auch zusätzliche Kontrollen. Der IESBA nennt das Hinzuziehen eines geeigneten Reviewers, der die Arbeit überprüft (IESBA, 2018, S. 160). Zudem wird ergebnisabhängige Vergütung untersagt („R905.7 Eine Praxis darf weder mittelbar noch

unmittelbar eine ergebnisabhängige Vergütung für eine Prüfung in Rechnung stellen.“, IESBA, 2018, S. 243).

Der Prüfungsausschuss soll die Sorgfalt der Arbeit im Mandat beurteilen (Hönsch & Kaspar, 2021, S. 80). Insgesamt zeigt sich damit ein Rahmen, der Qualität zugleich als Pflichterfüllung, als Unabhängigkeitszustand und als Systemleistung aus Prozessen und Kontrollen versteht.

4. Zivilrechtliche Haftung und Regulierungsmechanismen

4.1 Haftungsarchitektur in Deutschland

Ausgangspunkt der zivilrechtlichen Haftung des Abschlussprüfers ist der gesetzliche Pflichtmaßstab, der zugleich als Mindestanker für Prüfungsqualität fungiert. Der Abschlussprüfer ist zur „gewissenhaften und unparteiischen Prüfung“ verpflichtet (§ 323 Abs. 1 S. 1 HGB) (Deutscher Bundestag, Wissenschaftliche Dienste, 2020, S. 5). Zivilrechtliche Haftung knüpft daran an, wenn Pflichten verletzt werden.

Der Haftungstatbestand erfasst vorsätzliche und fahrlässige Pflichtverletzungen. Maßgeblich ist, dass bei Pflichtverletzung ein Schadensersatzanspruch entstehen kann, weil „wer vorsätzlich oder fahrlässig seine Pflichten verletzt, … zum Ersatz … verpflichtet“ ist (§ 323 Abs. 1 S. 3 HGB) (Wissenschaftliche Dienste des Deutschen Bundestages, 2021, S. 4). Damit ist die Haftung nicht auf Vorsatz beschränkt, sondern setzt bereits bei Fahrlässigkeit an.

Für Fahrlässigkeit sieht das deutsche Recht zugleich eine summenmäßige Haftungsbegrenzung vor. In der Darstellung der Wissenschaftlichen Dienste ist die Haftung „bei Fahrlässigkeit auf … 1 Million Euro … [bzw.] auf 4 Millionen Euro pro Prüfung begrenzt“ (Deutscher Bundestag, Wissenschaftliche Dienste, 2020, S. 5). Diese Obergrenzen bilden in der Literatur häufig die zentrale „Behandlungsvariable“, wenn Anreizwirkungen von Haftung diskutiert werden.

Die Haftungsbegrenzung wird als strukturelles, historisch stabiles Element beschrieben. Kuhmann (2025, S. 21) betont, dass gerade „die summenmäßige Begrenzung der Haftung des Abschlussprüfers“ Gesetzesreformen überdauert hat und im Zivilrecht als ungewöhnlich hervorgehoben wird. Für empirische Arbeiten ist dieser Befund relevant, weil er auf eine relativ konstante institutionelle Ausgangslage hindeutet.

Wichtig ist zudem die Frage, wer überhaupt Anspruchsinhaber ist. Nach der Darstellung der Wissenschaftlichen Dienste ist die Haftung im Grundsatz auf die geprüfte Kapitalgesellschaft als Auftraggeber beschränkt; Dritte können sich auf die Haftungsnormen „grundsätzlich nicht berufen“ (Deutscher Bundestag, Wissenschaftliche Dienste, 2020, S. 8). Entsprechend ordnet

Joha (2018) ein, dass einer Dritthaftung in Deutschland regelmäßig nur geringe Bedeutung zukomme (Joha, 2018, S. 95).

Formal existieren daneben deliktsrechtliche Anspruchsgrundlagen, die nicht unter die Haftungsobergrenze fallen. Gleichzeitig wird betont, dass diese deliktischen Ansprüche in der Praxis kaum Bedeutung erlangen (Deutscher Bundestag, Wissenschaftliche Dienste, 2020, S. 6). Für geschädigte Anleger verbleibt damit faktisch vor allem § 826 BGB als Option, der aber wegen hoher Hürden als seltene Durchsetzungsroute erscheint (Deutscher Bundestag, Wissenschaftliche Dienste, 2020, S. 10).

Die Architektur wird schließlich durch Durchsetzungs- und Versicherbarkeitsaspekte geprägt. In einer Anhörung wird die Versicherbarkeit als Grenze benannt, weil unbegrenzte Haftung praktisch nicht versicherbar sei (Deutscher Bundestag, Finanzausschuss, 2021, S. 13). Vor diesem Hintergrund wird deutlich, dass das deutsche Haftungsregime aus einer Kombination aus Sorgfaltspflicht, Fahrlässigkeitshaftung mit Cap sowie zugleich begrenzter Drittadressierung besteht, wodurch sich die Haftungsintensität als eher moderat einordnen lässt (Joha, 2018, S. 96).

4.2 Wirkungsmechanismen: Haftung - Prüfungsanstrengung, Honorar, Mandatsannahme, Dokumentation

Zivilrechtliche Haftung soll nach der Reformlogik als Hebel für Prüfungsqualität wirken, indem sie Unabhängigkeit und Sorgfalt ökonomisch absichert (IDW, 2021, S. 5). Für die Literatursynthese sind vier Pfade zentral: Prüfungsanstrengung, Honorarbildung, Mandatsannahme sowie Dokumentation und Qualitätsmanagement.

Erstens kann ein höheres Klagerisiko die Prüfungsanstrengung erhöhen. Pfaff und Ruffing-Straube beschreiben Effort als Funktion unter anderem von Litigation Risk und Reputationsrisiko; beides erhöht die erwarteten Kosten einer Prüfungsfehlerleistung und kann Skepsis und Prüfungstiefe stärken (Pfaff & Ruffing-Straube, 2021, S. 240). Trotz Haftungsdeckel kann ein Prüfungsfehler über Reputationsschäden Mandate kosten (Downar et al., 2021, S. 3). Dieser Pfad passt zur Output-Definition von Prüfungsqualität als Fehlerentdeckung und -berichterstattung, weil mehr Effort die Entdeckungswahrscheinlichkeit erhöhen soll.

Gleichzeitig ist offen, ob Haftungsdruck tatsächlich in mehr Arbeit statt in andere Reaktionen übersetzt wird. Pfaff und Ruffing-Straube berichten, dass Forschung eher Honorarreaktionen als Effort-Anpassungen findet, was einen direkten Qualitätseffekt abschwächen kann (Pfaff &

Ruffing-Straube, 2021, S. 241). Sethe warnt zudem, dass überproportionaler Haftungsdruck zu einer Verlagerung auf Haftungsvermeidung führen kann und damit die Prüfungsleistung schlechter werden kann (Sethe, 2012, S. 128).

Zweitens wirkt Haftung über Preise und Ressourcen. Audit Fees enthalten neben Aufwand auch Litigation Risk und Normalprofit; höhere Haftungsrisiken können daher als Risikoprämie in Honoraren erscheinen (Choi et al., 2010, S. 2). Völker betont, dass Honorarsteigerungen nicht nur Mehrarbeit, sondern auch eingepreiste Risiko- und Reputationsprämien spiegeln und deshalb nicht automatisch als Qualitätszuwachs zu lesen sind (Völker, 2021, S. 20). Brähler et al. ordnen Haftungsrisiko, erwartete Haftungsleistungen und Reputationsschäden ebenfalls als Bestandteile der Honorarlogik ein (Brähler et al., 2011, S. 3).

Drittens beeinflusst Haftung die Mandatsannahme und damit Marktstruktur. Wenn Prozessrisiken existenzbedrohend werden, können Prüfungsgesellschaften Risiken reduzieren, Prämien erhöhen oder sich aus riskanten Mandaten zurückziehen (Pfaff & Ruffing-Straube, 2021, S. 240). Berufsständische Stellungnahmen erwarten, dass Haftungsverschärfungen die Konzentration im Prüfungsmarkt verstärken und kleinere Praxen faktisch ausschließen können; im Extremfall finden Unternehmen dann keinen gesetzlichen Prüfer mehr (IDW, 2021, S. 5). Zusätzlich können steigende Versicherungsprämien kleinere Anbieter verdrängen und die Nachwuchsgewinnung erschweren (IDW, 2021, S. 15–16). Sethe verortet eine solche Eintrittsbarriere ausdrücklich in der fehlenden Versicherbarkeit des Haftungsrisikos (Sethe, 2012, S. 122). Als Gestaltungsansatz nennt Sethe eine Haftungsbegrenzung als Vielfaches des Honorars mit Selbstbehalt, um Anreize zu erhalten und Versicherbarkeit zu sichern (Sethe, 2012, S. 151).

Viertens wirkt Haftung auf Dokumentation und interne Systeme. Im Qualitätsmanagement können gezielt Qualitätsziele gesetzt werden, um Haftungsrisiken zu begrenzen, und die Mandatsannahme verlangt eine Analyse von Haftungs- und Reputationsrisiken (IDW, 2026, S. 10, 22). Auch die lange Aufbewahrungsfrist von Prüfungsdokumentation dient der Absicherung und erhöht den Nachweis- und Disziplinierungsdruck im Prüfungsprozess (IDW, 2026, S. 54). Insgesamt wird damit deutlich, dass Haftung Prüfungsqualität nicht nur über mehr Effort, sondern ebenso über Preis-, Selektions- und Dokumentationsmechanismen beeinflusst, wobei Nebenwirkungen wie Konzentration die intendierten Effekte überlagern können.

4.3 Ländervergleich/Institutionenvergleich als Kontrastfolie

Die in 4.1 skizzierte Haftungsarchitektur ist ein Ausgangspunkt. Für die Wirkungskette „Haftungsregime → Anreize → Prüfungsqualität“ ist entscheidend, wie stark ein Regime durchgesetzt wird und welche Adressaten Ansprüche erheben können (Langenbucher et al., 2020, S. 14–16; Günther, 2020, S. 79).

Als Kontrast verweist der Vergleich mit Frankreich auf schärferes Exponierungsniveau. Dort ist die Haftung verschuldensabhängig, zugleich aber grundsätzlich unbegrenzt und persönlich ausgestaltet (Wissenschaftliche Dienste des Deutschen Bundestages, 2021, S. 6). Hinzu kommt, dass eine gesamtschuldnerische Haftung möglich ist, wodurch sich das Risiko eines einzelnen Prüfers im Schadensfall erhöhen kann (Wissenschaftliche Dienste des Deutschen Bundestages, 2021, S. 6).

Das Vereinigte Königreich steht für einen anderen Typus: weniger gesetzliche Deckelung, mehr Vertragsgestaltung. Nach den Wissenschaftlichen Diensten bestimmt der Grundsatz der Vertragsfreiheit die Haftung über den Prüfungsvertrag. Zwar ist die Haftung grundsätzlich unbegrenzt, sie kann aber vertraglich begrenzt werden, sofern die Begrenzung „angemessen und vernünftig“ ist (Wissenschaftliche Dienste des Deutschen Bundestages, 2021, S. 7-8). Damit hängt der Haftungsdruck von Mandats- und Verhandlungssituationen ab und weniger von starren gesetzlichen Schwellen.

Für Deutschland prägen nicht allein gesetzliche Haftungsobergrenzen, sondern auch die praktische Durchsetzung. Langenbucher et al. beschreiben die Haftung in Deutschland als selten gerichtlich ausgetragen und halten fest, dass private enforcement fehlt (Langenbucher et al., 2020, S. 16). Konsistent damit ordnet Günther Deutschland im „Liability Standard Index“ als das Land ein, in dem eine Klage gegen Abschlussprüfer am schwierigsten ist (Günther, 2020, S. 79).

Diese Durchsetzung ist für die Interpretation von Haftungscaps wichtig. International seien Haftungsobergrenzen verbreitet; für Deutschland nennen Langenbucher et al. einen Maximalbetrag von 4 Mio. Euro. Zugleich argumentieren sie, dass höhere Grenzen die ex-ante-Anreize für skeptische und unabhängige Prüfungen stärken könnten (Langenbucher et al., 2020, S. 14; 17). Downar et al. ergänzen, dass selbst bei begrenzter Haftung das Reputationsrisiko im Fall eines Prüfungsversagens erheblich sein kann und Mandatsverluste auslösen kann (Downar et al., 2021, S. 3).

Im deutschen Reformdiskurs wird damit ein Abwägungsproblem sichtbar. Das IDW warnt, dass Haftungsverschärfungen Konsequenzen für die Marktkonzentration berücksichtigen müssen

(Institut der Wirtschaftsprüfer in Deutschland e. V. [IDW], 2021, S. 5). Im Finanzausschuss wird zudem die Idee variabler, proportionaler Haftungsobergrenzen diskutiert, weil fixe Grenzen als nicht anreizkompatibel gelten (Deutscher Bundestag, Finanzausschuss, 2021, S. 16). Vor diesem Hintergrund wird deutlich, dass länderübergreifende Evidenz zu Haftung und Prüfungsqualität nur dann wirklich gut vergleichbar ist, wenn Haftungsniveau, Vertragsgestaltung, Reputation und Durchsetzung gemeinsam als Kontextfaktoren mitgedacht werden.

5. Systematischer Literaturüberblick

5.1 Such- und Auswahlstrategie

Die Literaturrecherche wurde so angelegt, dass sie sowohl wissenschaftliche Evidenz als auch den aktuellen Reform- und Berufsstandsdiskurs abbildet. Ausgangspunkt war der deutsche Reformkontext, in dem Unabhängigkeits- und Haftungsfragen ausdrücklich mit dem Ziel einer Qualitätssteigerung verknüpft werden (Institut der Wirtschaftsprüfer in Deutschland e. V. [IDW], 2021, S. 5).

Recherchiert wurde in interdisziplinären Datenbanken für Wirtschafts- und Rechtswissenschaften (u. a. Google Scholar, Web of Science, EconBiz, SSRN). Ergänzend wurden Schneeballsuchen eingesetzt: Referenzlisten zentraler Arbeiten wurden rückwärts durchsucht und Zitationen vorwärts nachverfolgt, um einseitige Treffer durch einzelne Suchalgorithmen zu vermeiden.

Als Suchbegriffe wurden deutsche und englische Kombinationen genutzt, die den Wirkungszusammenhang „Haftungsregime–Anreiz–Prüfungsqualität“ abdecken. Zentral waren „auditor liability“, „litigation risk“, „audit quality“, „audit fees“, „reputation risk“, „enforcement“, „audit market concentration“, „mandatory rotation“, „non-audit services“, „fee cap“ und „joint audit“ sowie die entsprechenden deutschen Begriffe (z. B. Abschlussprüferhaftung, Prüfungsqualität, Haftungshöchstsumme). Qualitätsmaße wurden über Begriffe wie „accruals“ und „going concern modification“ gezielt mitgesucht.

Das Zeitfenster wurde grundsätzlich auf 2010 bis 2026 gelegt, um die Evidenz rund um neuere Reformwellen und aktuelle Qualitätsmanagementanforderungen abzudecken. Quellen, die die EU-Auditreform (Rotation und Beschränkung von Nichtprüfungsleistungen) als institutionellen Rahmen diskutieren, wurden unabhängig vom genauen Erscheinungsjahr aufgenommen (Maca & Tirdea, 2025, S. 246). Internationale Studien wurden einbezogen, wenn sie Regulierungsinstrumente wie Pflichtrotation als Qualitätsmaßnahme empirisch untersuchen

(Martani et al., 2021, S. 4) oder Joint Audits als policybasiertes Qualitätsinstrument einordnen (García-Blandón et al., 2021, S. 1379).

Aufgenommen wurden (1) peer-reviewte empirische Studien, (2) theorie- oder rechtswissenschaftliche Arbeiten, (3) Policy- und Positionspapiere sowie Gesetzesmaterialien, sofern sie Haftung oder verwandte Regulierungsmechanismen explizit mit Prüfungsqualität oder deren Proxies verbinden. Ausgeschlossen wurden Beiträge ohne klaren Bezug zu Haftung/Enforcement oder ohne operationalisierbares Qualitätsverständnis sowie rein beschreibende Marktberichte ohne Mechanismenbezug.

Die Treffer wurden zunächst über Titel und Abstract gesichtet, anschließend im Volltext geprüft und bei Mehrfachpublikationen wurde die jeweils aktuellere Fassung priorisiert. Der endgültige Kernkorpus umfasst zehn Quellen, die entweder Haftung direkt variieren oder Regulierungspakete (Rotation, NAS) mit Qualitätsproxies verbinden und damit im weiteren Verlauf für die Evidenzblöcke 5.3–5.5 anschlussfähig sind. Damit basiert die folgende Synthese auf einem bewusst gemischten, aber anhand klarer Kriterien selektierten Korpus, der die spätere Kodierung und die kritische Bewertung der Befunde unmittelbar unterstützt.

5.2 Kodierschema/Literaturmatrix: Variablen, Qualitätsmaße, Haftungsmaß, Kontext

Die in 5.1 ausgewählten Quellen werden in einer einheitlichen Literaturmatrix erfasst, damit die Befunde zur Frage „Haftungsregime – Anreize – Prüfungsqualität“ später vergleichbar sind. Ziel ist nicht, Einzelstudien nachzuerzählen, sondern ihre zentralen Aussagen entlang gleicher Kategorien zu kodieren und damit die Evidenzblöcke in 5.3–5.5 sowie die methodische Kritik in 6.1 vorzubereiten.

Erfasst werden zunächst formale Metadaten (Autor, Jahr, Land/Region, Untersuchungszeitraum) und die Evidenzart. Hier wird zwischen empirischer Forschung, konzeptioneller Analyse und policy-/positionsorientierten Texten unterschieden, weil sich daraus unterschiedliche Beweisstärken und Interessenlagen ergeben (IDW, 2021, S. 5).

Als „Behandlungsvariable“ wird die haftungs- bzw. regulierungsbezogene Größe kodiert. Dazu zählen Haftungsbegrenzungen und deren Höhe, Hinweise auf private Durchsetzung und Litigation-Kultur sowie ergänzende Regulierungsinstrumente wie Rotation oder Restriktionen bei Nichtprüfungsleistungen (Langenbucher et al., 2020, S. 14–17; Maca & Tirdea, 2025, S. 246). Zusätzlich wird erfasst, ob eine Quelle Haftung als monetären Sanktionsmechanismus

oder als Reputationsmechanismus versteht, da Reputationsverluste auch bei begrenzter Haftung erheblich sein können (Downar et al., 2021, S. 3).

Auf der Ergebnis-Seite wird das jeweilige Qualitätsmaß dokumentiert. Kodiert werden u. a. accruals-basierte Proxies und earnings-benchmarks, Going-Concern-Modifikationen, Prüfungshonorare sowie Haftungsfälle als seltene, aber präzise Surrogate (García-Blandón et al., 2021, S. 1393; Günther, 2020, S. 1, 53). Die Matrix hält fest, ob die Studie selbst auf Messlücken oder Proxy-Sensitivität hinweist (García-Blandón et al., 2021, S. 1393).

Für die Wirklogik werden die vermuteten Kanäle codiert: Anpassung der Prüfungsanstrengung an Litigation Risk, Preisreaktionen über Fees, ökonomische Bindungen über (abnorm) hohe Honorare sowie Selektion und Markteintrittsbarrieren (Pfaff & Ruffing-Straube, 2021, S. 240–241; Choi et al., 2010, S. 2; Joha, 2018, S. 148; Sethe, 2012, S. 122).

Zusätzlich wird je Quelle die Richtung des Effekts (+/0/–) kodiert. Bei Honoraren wird festgehalten, dass Fees Aufwand und Litigation Risk enthalten und Risikoprämien den Qualitätsrückschluss erschweren (Choi et al., 2010, S. 2; Joha, 2018, S. 148).

Schließlich werden Kontextmoderatoren festgehalten, etwa Stärke von Rechtsinstitutionen und Aufsicht, Marktsegment (Big 4 vs. Non-Big-4) und Bündelwirkungen mehrerer Reformen (Völker, 2021, S. 108; Martani et al., 2021, S. 3–5; Deutsches Aktieninstitut, 2021, S. 16). Insgesamt wird damit sichtbar, unter welchen Bedingungen Haftung eher als Qualitätsanreiz wirkt und wann Neben- und Messprobleme die Interpretation dominieren.

5.3 Evidenzblock A: Befunde „Haftung erhöht Qualität“

Die im Evidenzblock A gebündelten Quellen argumentieren, dass ein höheres Haftungs- bzw. Sanktionsrisiko die Abschlussprüfung über Anreize und Kontrollmechanismen verbessern kann. In der Literaturmatrix werden solche Aussagen als positiver Qualitätseffekt kodiert, wenn sie sich auf Effort, Skepsis, Unabhängigkeit oder direkte Qualitätsproxies beziehen und gleichzeitig Kontextbedingungen benennen. Für die Kodierung wird daher jeweils notiert, welches Qualitätsmaß konkret genutzt wird und ob es Output oder Prozess abbildet.

Auf der Mechanismenseite setzen Pfaff und Ruffing-Straube (2021, S. 240) an der Prüfungsanstrengung an. Sie stellen heraus, dass Effort nicht nur vom Berufsethos, sondern auch vom Klage- und Reputationsrisiko abhängt. Steigt das Litigation Risk, ist damit konzeptionell ein höherer Prüfungsaufwand plausibel, der sich in Prozessqualität (z. B. kritischere Prüfungshandlungen) und letztlich Outputqualität (Entdeckung und Berichterstattung von Fehlern) übersetzen kann.

Günther (2020, S. 77) stützt diese Logik über die Präventivfunktion zivilrechtlicher Haftung. Die Möglichkeit, für fehlerhafte Prüfungen in Anspruch genommen zu werden, kann ex ante dazu führen, dass Prüfungen sorgfältiger geplant und dokumentiert werden, weil die erwarteten Kosten eines Prüfungsfehlers steigen. In der Kodierung wird dies als Anreizkanal erfasst, auch wenn die empirische Messung des tatsächlichen Effort häufig indirekt bleibt.

Empirisch wird der positive Haftungs-Qualitäts-Zusammenhang in Völker (2021) vor allem als institutionsabhängig beschrieben. In Regionen mit starken Rechtsinstitutionen und entwickelten Märkten sind Haftungs- und Reputationsrisiken höher, weshalb Prüfungspartner die Prüfungsqualität grundsätzlich höher ansetzen (Völker, 2021, S. 108). Zudem berichtet Völker (2021, S. 122) einen Befund, wonach eine Erhöhung des Haftungsrisikos mit einer höheren Prüfungsqualität einhergeht; in der Literaturmatrix wird dabei der Kontext „starke Institutionen“ als Moderator festgehalten.

Als verwandtes Accountability-Instrument wird Rotation diskutiert. Martani et al. (2021, S. 3) finden für Indonesien einen positiven Zusammenhang zwischen verpflichtender Prüfungsfirmarotation und Prüfungsqualität. Der Effekt ist jedoch heterogen: Bei Non-Big-4-Prüfern wirkt Firm-Rotation signifikant qualitätssteigernd, während bei Big-4-Prüfern Partner-Rotation als ausreichend beschrieben wird (Martani et al., 2021, S. 14). Für die Synthese ist relevant, dass regulatorische Eingriffe somit je nach Marktsegment unterschiedliche Qualitätsgewinne auslösen können.

Auch policybasierte Quellen liefern pro-Argumente, die als normative Evidenz zu kodieren sind. Im Finanzausschuss wird etwa vertreten, höhere Haftungssummen seien sinnvoll, weil dadurch Anreize für eine sorgfältigere Prüfung gesetzt würden (Deutscher Bundestag, Finanzausschuss, 2021, S. 7). Langenbucher et al. (2020, S. 17) empfehlen ebenfalls, bestehende Haftungsobergrenzen deutlich anzuheben, um ex ante skeptische und unabhängige Audits zu fördern. Insgesamt wird damit deutlich, dass positive Qualitätseffekte in diesem Block vor allem dann erwartet oder beobachtet werden, wenn Haftungsdruck durchsetzbar ist und in ein institutionell tragfähiges Umfeld eingebettet bleibt.

5.4 Evidenzblock B: Befunde „Haftung erhöht Qualität nicht / Nebenwirkungen dominieren“

Die im Evidenzblock B zusammengeführten Quellen stellen den erwarteten Qualitätsgewinn einer Haftungsverschärfung in Frage oder rücken Nebenwirkungen in den Vordergrund. Auffällig ist, dass ein Teil der Argumente aus berufsständischen und policy-nahen

Stellungnahmen stammt und damit eher Wirkungsbehauptungen als kausal identifizierte Effekte liefert.

Für den deutschen Reformdiskurs formuliert das IDW ausdrücklich die Erwartung, dass eine deutliche Verschärfung der Wirtschaftsprüferhaftung die Prüfungsqualität nicht erhöht (Institut der Wirtschaftsprüfer in Deutschland e. V. [IDW], 2021, S. 5). Als Begründung wird nicht primär der fehlende Anreizmechanismus bestritten, sondern die Gefahr, dass die Belastung die Leistungsfähigkeit und die Marktstruktur des Prüfungsmarkts beeinträchtigt.

Im gleichen Dokument wird vor einer Ausweitung hin zu unbegrenzter und praktisch nicht versicherbarer Haftung gewarnt (IDW, 2021, S. 1). Zudem wird argumentiert, dass strengere Haftung die Konzentration verstärkt und kleinere sowie mittelständische Prüfungspraxen von Abschlussprüfungen verdrängen kann (IDW, 2021, S. 5). Daraus folgt im Extrem ein Versorgungsproblem, wenn Unternehmen in schwieriger Lage keinen Prüfer finden, sowie ein langfristiger Qualitätsdruck über erschwerte Nachwuchsgewinnung (IDW, 2021, S. 5, 16).

Einen eher verhaltensökonomischen Gegenkanal diskutieren Pfaff und Ruffing-Straube: Zwar hängt Prüfungsanstrengung konzeptionell vom Litigation Risk ab, empirische Forschung deute jedoch darauf hin, dass Prüfungsgesellschaften höhere Risiken eher über steigende Honorare als über zusätzliche Prüfungsanstrengung verarbeiten (Pfaff & Ruffing-Straube, 2021, S. 241). In dieser Lesart steigt der Preis der Prüfung, ohne dass der Output zwingend besser wird.

Deutsches Aktieninstitut verlagert die Kritik auf das Regulierungsbündel und erwartet im Zusammenspiel von Haftungsverschärfung, Rotation und Einschränkungen von Nichtprüfungsleistungen eine weitere Marktkonzentration, die die Prüfungsqualität nicht verbessert (Deutsches Aktieninstitut, 2021, S. 6). Es prognostiziert außerdem Rückzüge von Prüfungsgesellschaften aus der Abschlussprüfung und negative Effekte auf die Berufswahl, was den Qualitätswettbewerb zusätzlich schwächen kann (Deutsches Aktieninstitut, 2021, S. 17).

Sethe betont die Versicherbarkeit als Engpass: Haftungsrecht könne als Eintrittsbarriere wirken, weil kleinere Anbieter das Risiko nicht versichern können (Sethe, 2012, S. 122). Zu hohe Haftungsdrohungen können zudem zu defensiven Strategien führen, bei denen Haftungsvermeidung die eigentliche Prüfungsaufgabe überlagert (Sethe, 2012, S. 128). Haucap et al. ergänzen, dass höhere Haftungssummen gerade im PIE-Segment Wettbewerbsrisiken erzeugen und kleinere Anbieter von Bewerbungen abhalten können (Haucap et al., 2022, S. 30). Unter diesen Bedingungen kann teilweise der intendierte Qualitätsanreiz durch Versicherungsengpässe und Marktaustritte überlagert werden (Sethe, 2012, S. 122). Damit verdichtet sich ein konsistentes Nebenwirkungsnarrativ: Haftung kann zwar Anreize setzen,

wird aber als Instrument dann fragwürdig, wenn sie vor allem Kosten, Selektion und Konzentration treibt und damit die Bedingungen für Prüfungsqualität verschlechtert.

5.5 Evidenzblock C: Interdependenzen und gemischte Effekte

Mehrere Beiträge betrachten Haftung nicht isoliert, sondern als Teil eines Regulierungsbündels, das Unabhängigkeit, Marktstruktur und wahrgenommene Prüfungsqualität zugleich adressiert. Für die Literatursynthese bedeutet das, dass beobachtete Qualitätseffekte häufig aus der Kombination von Haftungsregeln, Rotation und Beschränkungen der Nichtprüfungsleistungen entstehen und nicht eindeutig einem einzelnen Instrument zugeordnet werden können (Deutsches Aktieninstitut, 2021, S. 16; Maca & Tirdea, 2025, S. 246).

Im Kontext der EU-Auditreform werden insbesondere verpflichtende Prüferwechsel sowie Restriktionen für Non-Audit Services als zentrale Hebel genannt (Maca & Tirdea, 2025, S. 246). Maca und Tirdea berichten, dass Investoren auf die Einschränkung von Nichtprüfungsleistungen und eine Kappung der Gesamthonorare positiv reagieren, was als Indiz für gestiegenes Vertrauen in die Abschlussprüfung interpretiert werden kann (Maca & Tirdea, 2025, S. 249). Gleichzeitig weisen sie darauf hin, dass die Evidenz zur Wirkung der Pflichtrotation auf Unabhängigkeit, Mandantenbeziehung und Konzentration insgesamt knapp bleibt, sodass die Richtung des Nettoeffekts offen ist (Maca & Tirdea, 2025, S. 249).

Eine ähnliche Ambivalenz zeigt sich bei Joint Audits. García-Blandón et al. leiten zwar konzeptionell einen potenziellen Qualitätsgewinn aus dem „four eyes“-Gedanken ab, stellen diesem aber Free-Riding-Anreize gegenüber, die zu geringeren Prüfungsressourcen führen können (García-Blandón et al., 2021, S. 1380–1381). Empirisch finden sie bei Verwendung diskretionärer Accruals keinen Unterschied zwischen Joint- und Single-Audit-Regimen, während ein Benchmark-Maß (Just-beating) Hinweise auf niedrigere Qualität in Frankreich liefert (García-Blandón et al., 2021, S. 1378, 1384). Damit wird sichtbar, dass gemischte Befunde teils aus der Wahl des Qualitätsproxys resultieren.

Interdependenzen entstehen zudem durch funktionale Substitution: Bei begrenzter zivilrechtlicher Haftung kann Reputationsrisiko als disziplinierender Mechanismus wirken. Downar et al. betonen, dass trotz Haftungsbegrenzung erhebliche Reputationsverluste und Mandatsverluste drohen, wenn eine Prüfung scheitert (Downar et al., 2021, S. 3). In solchen Umfeldern ist es plausibel, dass zusätzliche Haftungsschärfen weniger „neue“ Anreize setzen als in Systemen mit schwächerer Reputation. Schließlich verweist die Literatur auf Bündelnebenwirkungen über Marktstruktur.

Die positive Marktreaktion auf NAS-Beschränkungen und Fee-Caps spricht dafür, dass schon die Wahrnehmung von Unabhängigkeit als Qualitätsdimension direkt wirkt (Maca & Tirdea, 2025, S. 249). Vor diesem Hintergrund zeigt sich, dass die Frage nach Haftung und Prüfungsqualität praktisch nur als Abwägung innerhalb eines Regelungsportfolios beantwortet werden kann, in dem Vertrauen, messbare Qualität und Wettbewerbsfolgen gleichzeitig zu berücksichtigen sind (Maca & Tirdea, 2025, S. 249; Deutsches Aktieninstitut, 2021, S. 16).

6. Kritische Würdigung

6.1 Kritische Diskussion der Literatur

Die in Kapitel 5 gebündelte Literatur liefert ein uneinheitliches Bild darüber, ob zivilrechtliche Haftung die Prüfungsqualität erhöht oder ob Nebenwirkungen überwiegen. Als plausibler Ausgangspunkt gilt der Anreizkanal, dass Prüfungsanstrengung sowohl vom Klagerisiko als auch vom Reputationsrisiko beeinflusst wird (Pfaff & Ruffing-Straube, 2021, S. 240).

Zur Konsistenz: Ein Teil der Beiträge und Empfehlungen erwartet höhere Qualität durch stärkere monetäre Sanktionen, etwa über deutlich erhöhte Haftungslimits (Langenbucher et al., 2020, S. 17) oder über den Befund, dass stärkerer Haftungsdruck mit höherer Qualität einhergehen kann (Völker, 2021, S. 122). Dem stehen berufsständische Stellungnahmen entgegen, die eine Qualitätswirkung verneinen und stattdessen Verdrängungseffekte betonen (IDW, 2021, S. 5).

Diese Divergenz ist auch kausal schwer zu trennen, weil Haftung selten isoliert durch einzelne politische Reformen verändert wird. So wird argumentiert, dass Haftung im Zusammenspiel mit Rotation und Einschränkungen von Nichtprüfungsleistungen die Marktstruktur weiter konzentriert, ohne die Qualität zu verbessern (Deutsches Aktieninstitut, 2021, S. 6).

Für die Kausalität ist zudem entscheidend, ob Haftung tatsächlich in mehr Skepsis und mehr Prüfungsarbeit übersetzt wird oder nur in Preise. Pfaff und Ruffing-Straube berichten, dass Prüfungsgesellschaften höhere Risiken eher über höhere Honorare als über höhere Anstrengungen verarbeiten (Pfaff & Ruffing-Straube, 2021, S. 241), was einen reinen „Effort"-Effekt empirisch schwer beobachtbar macht.

Messprobleme entstehen, weil zentrale Proxies mehrdeutig sind. Audit Fees enthalten Aufwand, Litigation Risk und Normalgewinne (Choi et al., 2010, S. 2) und können bei ökonomischer Abhängigkeit mit niedrigerer Qualität zusammenhängen (Choi et al., 2010, S. 2–3). Zudem ist unklar, ob Haftungsrisiken in Preisverhandlungen überhaupt sauber eingepreist werden (Brähler et al., 2011, S. 6).

Auch accrual-basierte Qualitätsmaße liefern kein stabiles Bild, weil unterschiedliche Proxies zu unterschiedlichen Ergebnissen führen. García-Blandón et al. zeigen eine deutliche Proxy-Lücke und berichten, dass Joint Audits je nach Maß (Accruals vs. Benchmark-Indikatoren) keinen oder sogar negativen Qualitätseffekt nahelegen (García-Blandón et al., 2021, S. 1378, 1384, 1393).

Endogenität verschärft das Problem, wenn Qualitätsentscheidungen selbst Klagerisiken verändern. Günther zeigt, dass bei Going-Concern-Modifikationen die Berücksichtigung der Endogenität die Beziehung zum Klagerisiko substanziell verändert (Günther, 2020, S. 146).

Schließlich ist die Generalisierbarkeit stark institutionsabhängig. Begrenzte Haftung kann durch Reputationsverluste teilweise kompensiert werden (Downar et al., 2021, S. 3), während in Deutschland zugleich hohe Durchsetzungshürden und geringe private Rechtsdurchsetzung den Abschreckungseffekt schwächen können (Deutscher Bundestag, Wissenschaftliche Dienste, 2020, S. 10; Langenbucher et al., 2020, S. 16).

Vor diesem Hintergrund erscheint ein linearer Zusammenhang „mehr Haftung = mehr Qualität“ weder empirisch robust noch konzeptionell zwingend, zumal zu hohe und nicht versicherbare Risiken Markteintritt und Wettbewerb beeinträchtigen können (Sethe, 2012, S. 122; Haucap et al., 2022, S. 30). Insgesamt wird damit deutlich, dass die Literatur eher auf bedingte Effekte hinweist, bei denen Enforcement, Reputation und Marktstruktur darüber entscheiden, ob Haftung Qualitätsanreize setzt oder Qualität über defensive Strategien und Konzentration indirekt gefährdet.

6.2 Eigene Bewertung und Implikationen

Die in Kapitel 5 gebündelte Evidenz legt nahe, dass zivilrechtliche Haftung Prüfungsqualität nicht automatisch erhöht. Der Anreizkanal ist plausibel, aber er wird leicht durch Preisreaktionen, Risikoselektion und Marktstrukturfolgen überlagert. Vor allem dort, wo Durchsetzung schwach ist, bleibt Haftung ein Papierreiz und ersetzt nicht die tägliche Qualitätssteuerung.

Regulierung sollte daher weniger auf eine pauschale Verschärfung setzen, sondern auf ein Design, das Anreiz und Versicherbarkeit verbindet. Ein praktikabler Kompromiss ist eine Haftungsobergrenze als Vielfaches des Prüfungshonorars, ergänzt um einen Selbstbehalt, damit trotz Deckel ein spürbarer Sorgfaltsanreiz bleibt (Sethe, 2012, S. 151). In dieselbe Richtung zielen Vorschläge, Haftungsobergrenzen risikoorientiert und etwa honorarabhängig zu variieren (Deutscher Bundestag, Finanzausschuss, 2021, S. 12).

Gleichzeitig muss jede Erhöhung die Wettbewerbsseite mitdenken. In den Gesetzesmaterialien wird ausdrücklich davor gewarnt, dass insbesondere mittelständische Praxen bei zu hoher Haftung aus dem Markt gedrängt werden können (Deutscher Bundestag, Finanzausschuss, 2021, S. 30). Zudem wird eine weitere Konzentration des Prüfungsmarkts als mögliche Folge benannt, die bei der Ausgestaltung mitzuwägen ist (Deutscher Bundestag, Finanzausschuss, 2021, S. 44).

Umgekehrt ist auch ein zu niedriger Deckel riskant, weil er Sorgfaltspflichten entwerten und damit Qualität schwächen kann (Sethe, 2012, S. 136). Deshalb spricht vieles für eine Haftung, die nach Mandatsrisiko staffelt und zugleich versicherbar bleibt, statt entweder Symbolpolitik oder ruinöse Exposition zu erzeugen. Das erhöht die Planbarkeit für Beteiligte.

Für die Prüfungspraxis folgt, dass Haftungsdruck vor allem ex ante über Qualitätsmanagement wirkt. Das IDW betont, dass Praxen zusätzliche Qualitätsziele auch zur Begrenzung von Haftungsrisiken setzen können (Institut der Wirtschaftsprüfer in Deutschland e. V. [IDW], 2026, S. 10). Zentral sind Mandatsannahme- und Fortführungsprozesse, die Haftungs- und Reputationsrisiken systematisch prüfen (IDW, 2026, S. 22), sowie eine robuste Dokumentation mit langfristiger Aufbewahrung als Absicherung im Streitfall (IDW, 2026, S. 54).

Für die Forschung bleibt die Messfrage entscheidend. Honorare sind wegen eingepreister Haftungs- und Reputationsprämien nur begrenzt als Qualitätsproxy geeignet (Joha, 2018, S. 148). Zudem zeigen Proxy-vergleichende Arbeiten, dass sich Befunde je Messansatz deutlich unterscheiden können, wenn Qualität vor allem über Accruals gemessen wird (García-Blandón et al., 2021, S. 1393). Mehr Studien mit Zugang zu internen Aufwanddaten könnten den Effort-Kanal besser identifizieren (Günther, 2020, S. 270).

Damit Haftungsregeln tatsächlich Qualitätsanreize entfalten, braucht es zugleich klare, einklagbare Pflichten zur skeptischen Prüfung und eine Durchsetzung, die die Kosten von Verstößen real erhöht (Langenbucher et al., 2020, S. 17; Martani et al., 2021, S. 5), sodass insgesamt eine kalibrierte Haftungsarchitektur am ehesten geeignet erscheint, Qualität zu stützen, ohne den Markt funktionsunfähig zu machen.

7. Fazit und Ausblick

7.1 Beantwortung der Forschungsfrage(n) und Rückbezug auf Hypothesen

Die Arbeit fragte, wie zivilrechtliche Haftungsregelungen das Verhalten von Abschlussprüfern und damit die Prüfungsqualität beeinflussen. Die Literatur deutet auf mehrere, teils gegenläufige Wirkpfade hin.

H1 (Anreiz-/Effort-Hypothese) wird grundsätzlich gestützt, sofern Haftungs- und Reputationsrisiken tatsächlich wirksam sind. Prüfungsanstrengung hängt konzeptionell vom Litigation Risk und vom Reputationsrisiko ab (Pfaff & Ruffing-Straube, 2021, S. 240); in Umfeldern mit starken Rechtsinstitutionen wird deshalb eine höhere Qualität erwartet (Völker, 2021, S. 108). Entsprechend plädieren Reformvorschläge dafür, Haftungsobergrenzen deutlich anzuheben, um ex ante Skepsis und Unabhängigkeit zu stärken (Langenbucher et al., 2020, S. 17).

H2 (Kosten-/Nebenwirkungs-Hypothese) relativiert diesen positiven Kanal. Empirische und konzeptionelle Hinweise sprechen dafür, dass erhöhte Risiken eher in höheren Honoraren als in zusätzlichem Prüfungsaufwand verarbeitet werden (Pfaff & Ruffing-Straube, 2021, S. 241). Gleichzeitig können nicht versicherbare oder stark steigende Haftungsrisiken Marktaustritte und Konzentration fördern (Institut der Wirtschaftsprüfer in Deutschland e. V. [IDW], 2021, S. 5) und für kleinere Praxen existenzgefährdend sein (Haucap et al., 2022, S. 30), was mittelbar die Qualität gefährden kann.

H3 (Substitutionshypothese) wird als bedingte Aussage plausibel: Trotz begrenzter Haftung kann Reputationsverlust erhebliche Sanktionen auslösen (Downar et al., 2021, S. 3). Wo private Durchsetzung selten ist, fehlt jedoch ein Teil des Haftungsdrucks (Langenbucher et al., 2020, S. 16).

H4 (Interdependenzhypothese) wird durch Hinweise gestützt, dass Haftung im Reformbündel mit Rotation und NAS-Restriktionen auf Konzentration und Qualität wirkt (Deutsches Aktieninstitut, 2021, S. 6). Die Evidenz zu Rotation ist heterogen (Martani et al., 2021, S. 3), und Joint-Audit-Befunde hängen vom Qualitätsproxy ab (García-Blandón et al., 2021, S. 1378, 1384). Insgesamt spricht die Literatur damit eher für eine kalibrierte Haftungsarchitektur, deren Qualitätsnutzen von Enforcement, Reputation und Marktstrukturfolgen abhängt. Für Deutschland bedeutet dies, dass eine bloße Erhöhung von Haftungssummen ohne wirksame Durchsetzung und ohne Blick auf Versicherbarkeit kaum verlässlich zu mehr Qualität führt, während eine zu niedrige Haftung Anreize abschwächen kann im Ergebnis.

7.2 Limitationen der Arbeit

Diese Arbeit stützt sich vollständig auf eine literaturbasierte Evidenzsynthese. Damit hängen Reichweite und Präzision der Aussagen stark davon ab, wie die ausgewerteten Studien Prüfungsqualität überhaupt messbar machen und welche Daten in den jeweiligen Kontexten verfügbar sind.

Ein zentrales Problem ist die Proxy-Abhängigkeit. García-Blandón et al. (2021, S. 1393) zeigen, dass Befunde je nach Qualitätsmaß (z. B. discretionary accruals vs. earnings benchmarks) auseinanderlaufen und die Forschung Prüfungsqualität bislang häufig über Accruals operationalisiert. Zudem weisen dieselben Autoren darauf hin, dass nicht beobachtete Faktoren die berichteten Ergebnisse beeinflusst haben könnten, was die Kausalinterpretation im Ländervergleich begrenzt (García-Blandón et al., 2021, S. 1393).

Auch die Verwendung von Prüfungshonoraren ist ambivalent. Wenn Abschlussprüfer Risikoprämien für mögliche Haftungs- und Reputationsschäden einpreisen, wird ein Honorarunterschied nicht eindeutig als Qualitätsunterschied interpretierbar (Joha, 2018, S. 148).

Haftungsfälle erscheinen zwar als präziser Output-Indikator, sind empirisch jedoch selten. Dadurch sinken Aussagekraft und statistische Trennschärfe von Analysen, die auf beobachteten Haftungsfällen beruhen (Günther, 2020, S. 53).

Hinzu kommt, dass zentrale Output-Proxies selbst endogen sein können. Wird etwa die Endogenität zwischen Going-Concern-Modifikationen und Klagerisiko berücksichtigt, verändern sich die Effektaussagen deutlich (Günther, 2020, S. 146).

Schließlich ist die Übertragbarkeit institutionell eingeschränkt. Unterschiede in Aufsicht und Durchsetzung verändern die Kosten von Verstößen und damit die Wirkung von Haftung und Regulierung (Martani et al., 2021, S. 5), während die Evidenz zur Pflichtrotation insgesamt als knapp beschrieben wird (Maca & Tirdea, 2025, S. 249). Evidenzbasis bleibt knapp und uneinheitlich (Maca & Tirdea, 2025, S. 249). Vor diesem Hintergrund lassen sich die Ergebnisse nur als konditionale, kontextabhängige Aussagen lesen.

7.3 Ausblick

Für die Regulierung spricht die Literatur weniger für „mehr oder weniger Haftung“ als für ein besser kalibriertes Design. Im Finanzausschuss wird eine risikoorientierte, etwa am Prüfungshonorar ausgerichtete Haftungsobergrenze als effizientere Alternative zu fixen Grenzen diskutiert (Deutscher Bundestag, Finanzausschuss, 2021, S. 12, 16). Sethe (2012, S. 151) konkretisiert dies als Cap in einem Vielfachen des Honorars, ergänzt um einen Selbstbehalt, um Versicherbarkeit und Anreizwirkung zugleich zu adressieren.

Als Gegenmodell wird eine deutliche Anhebung bestehender Haftungsdeckel zur Stärkung ex-ante skeptischer und unabhängiger Prüfungen vorgeschlagen (Langenbucher et al., 2020, S. 17).

Jede Reform sollte dabei auf Nebenwirkungen wie Rückzug von Prüfungsgesellschaften aus der Abschlussprüfung achten (Deutsches Aktieninstitut, 2021, S. 17).

Die Wirkung neuer Haftungsregeln lässt sich nur prüfen, wenn begleitend messbare Kriterien festgelegt werden. Zentral sind Konzentrationsmaße und Marktanteile, weil die Ausgangslage in Deutschland bereits sehr hoch ist (Maca & Tirdea, 2025, S. 247) und Reformen teils keine Verschiebung zeigen (Maca & Tirdea, 2025, S. 251). Ergänzend sollten Versicherbarkeitsindikatoren wie Prämienentwicklung sowie Anbieterzahlen im Prüfungsmarkt beobachtet werden. Auch die Zahl der Mandatsniederlegungen und die Teilnahme am Qualitätskontrollsystem können als Frühwarnindikatoren für Angebotsverknappung dienen, wenn Haftungsrisiken steigen (Deutsches Aktieninstitut, 2021, S. 17).

Für die Forschung folgt daraus ein Bedarf an Daten, die den Prüfungsaufwand näher abbilden, statt ihn nur indirekt über Honorare zu schätzen (Günther, 2020, S. 270). Parallel kann die Praxis über Qualitätsmanagementziele, die Haftungsrisiken explizit einbeziehen, einheitlichere interne Messpunkte schaffen (IDW, 2026, S. 10), sodass künftige Evaluierungen Effort-, Marktstruktur- und Qualitätsindikatoren gemeinsam auswerten können. Das würde zugleich bessere Kausaltests erlauben.

Literaturverzeichnis

Behringer, S. (2019). Fraud in der prüferischen Praxis der Jahresabschlussprüfung: Die Suche nach dem faulen Apfel. WP Praxis, 5, 137–143.

Brähler, G., Brune, P., & Göttsche, M. (2011). Determinanten der Prüfungshonorare von Unternehmen aus der Medienbranche: Eine empirische Untersuchung unter Einsatz einer Regressionsanalyse. VERLAG proWiWi e.V. https://nbn-resolving.de/urn:nbn:de:gbv:ilm1-2011200497

Choi, J.-H., Kim, J.-B., & Zang, Y. (2010). Do abnormally high audit fees impair audit quality? Auditing: A Journal of Practice & Theory, 29(2), 115–140. https://doi.org/10.2308/aud.2010.29.2.115

Deutscher Bundestag, Finanzausschuss. (2021, March 15). Wortprotokoll der 121. Sitzung [Transcript].

Deutscher Bundestag, Wissenschaftliche Dienste. (2021, April 7). Einzelfragen zur Haftung von Wirtschaftsprüfern in ausgewählten europäischen Staaten (Sachstand, WD 7-3000-029/21). Deutscher Bundestag.

Deutscher Bundestag, Wissenschaftliche Dienste. (2020). Einzelfragen zur Haftung von Wirtschaftsprüfern (Sachstand, WD 7-3000-090/20). Author.

Downar, B., Ernstberger, J., & Koch, C. (2021). Who makes partner in Big 4 audit firms? Evidence from Germany. Accounting, Organizations and Society, 91, Article 101176. https://doi.org/10.1016/j.aos.2020.101176

García Blandón, J., Castillo-Merino, D., Argilés Bosch, J. M., & Ravenda, D. (2021). Mandatory joint audit and audit quality in the context of the European blue chips. Journal of Business Economics and Management, 22(5), 1378–1395. https://doi.org/10.3846/jbem.2021.14959

Günther, J. (2020). Die Unabhängigkeit des Abschlussprüfers bei privaten Unternehmen in Deutschland: Eine empirische Analyse im Kontext der Honorare für Prüfung und Beratung. Würzburg University Press. https://doi.org/10.25972/WUP-978-3-95826-117-4

Haucap, J., Kehder, C., & Prüfer, M. (2022). Funktionsdefizite auf dem Wirtschaftsprüfungsmarkt: Eine Studie im Auftrag der Mazars GmbH & Co. KG. DICE Consult GmbH.

Hönsch, H., & Kaspar, M. (2021). Der Prüfungsausschuss: Best Practices einer effizienten Überwachung (6., überarb. Aufl.). PricewaterhouseCoopers GmbH Wirtschaftsprüfungsgesellschaft.

Institut der Wirtschaftsprüfer in Deutschland e. V. (IDW). (2026). Entwurf eines IDW Qualitätsmanagementstandards: Anforderungen an das Qualitätsmanagement in der Wirtschaftsprüferpraxis (IDW EQMS 1 (02.2026)). IDW Verlag GmbH.

Institut der Wirtschaftsprüfer in Deutschland e. V. (IDW). (2021, October 28). Kommunikation von Prüfungsqualität: Vorschläge für einen strukturierten Dialog über Prüfungsqualität unter Berücksichtigung von Audit Quality Indicators (IDW Positionspapier). Institut der Wirtschaftsprüfer in Deutschland e. V.

International Ethics Standards Board for Accountants. (2018). Handbuch des internationalen Verhaltenskodex für Berufsangehörige: einschließlich internationaler Unabhängigkeitsstandards (Ausgabe 2018). International Federation of Accountants.

Joha, P. (2018). Prüfungshonorare und Prüferwahl großer, nicht börsennotierter Unternehmen. Würzburg University Press. https://nbn-resolving.org/urn:nbn:de:bvb:20-opus-146508

Kuhmann, L. (2025). Die Summenbegrenzung der Haftung von Abschlussprüfern (Abhandlungen zum Deutschen und Europäischen Gesellschafts- und Kapitalmarktrecht; Vol. 265). Duncker & Humblot.

Langenbucher, K., Leuz, C., Krahnen, J. P., & Pelizzon, L. (2020). What are the wider supervisory implications of the Wirecard case? (Study requested by the ECON committee, PE 651.385). Economic Governance Support Unit, Directorate-General for Internal Policies, European Parliament. http://www.europarl.europa.eu/supporting-analyses

Maca, A. I., & Tirdea (Predoaia), O. V. (2025). Audit market concentration for listed entities in the European Union and the impact following the latest audit reform. Proceedings of the 19th International Conference on Business Excellence, 246–256. https://doi.org/10.2478/picbe-2025-0022

Martani, D., Rahmah, N. A., Amarullah, F., & Anggraita, V. (2021). Impact of audit tenure and audit rotation on the audit quality: Big 4 vs non Big 4. Cogent Economics & Finance, 9(1), 1–19. https://doi.org/10.1080/23322039.2021.1901395

Pfaff, D., & Ruffing-Straube, P. (2021). Aufdeckung von Bilanzbetrug: Sind Leerverkäufer besser als Wirtschaftsprüfer? Die Unternehmung, 75(2), 238–247. https://doi.org/10.5771/0042-059X-2021-2-238

Sethe, R. (2012). Rechtspolitische Überlegungen zur Haftung der Revisionsstelle. In R. H. Weber & P. R. Isler (Eds.), Verantwortlichkeit im Unternehmensrecht VI (pp. 121–158). Schulthess Juristische Medien AG.

Völker, U. G. (2021). Die Prüfungshonorare branchenspezialisierter Wirtschaftsprüfer in der deutschen Versicherungsbranche. Würzburg University Press. https://doi.org/10.25972/WUP-978-3-95826-157-0

Anhang A: Kodierschema

Kodierdimension	Was wird erfasst?	Typische Ausprägungen / Codes	Ankerquelle(n)
Evidenzart	Welche Art von Quelle und Argumenttyp?	Empirische Studie; konzeptionell/theoretisch; Policy-/Positionspapier; Rechts-/Parlamentsmaterial	(IDW, 2021, S. 5)
Kontext/ Institutionen	Setting, das Haftungs- und Qualitätswirkungen moderiert	Land/Regime; Stärke Rechtsinstitutionen/Aufsicht; Marktsegment (Big 4 vs. Non-Big-4)	(Völker, 2021, S. 108); (Martani et al., 2021, S. 3–5)
Haftungs-/ Regulierungs-variable	Welche „Hebel“ werden analysiert?	Liability caps/Höhe; private enforcement/Litigation-Kultur; Rotation; NAS-Restriktionen; Fee Caps; Joint Audit; Reformbündel	(Langenbucher et al., 2020, S. 14–17); (Maca & Tirdea, 2025, S. 246)
Qualitätsmaß (Outcome)	Woran wird Prüfungsqualität festgemacht?	Accruals; Earnings Benchmarks; GCM; Audit Fees; Haftungsfälle	(García-Blandón et al., 2021, S. 1393); (Günther, 2020, S. 1, 53)
Wirkmechanismus (Kanal)	Wie soll Haftung/Regulierung wirken?	Effort/Skepsis; Fees/Risikoprämie; ökonomische Bindung (abnormal fees/quasi-rents); Selektion/Markteintritt; Reputation als Substitut	(Pfaff & Ruffing-Straube, 2021, S. 240–241); (Choi et al., 2010, S. 2–3, 7); (Downar et al., 2021, S. 3)
Proxy-/ Messprobleme	Welche Einschränkungen nennt die Quelle selbst?	Fees = Effort + Litigation Risk; Risikoprämien verzerren; Proxy-Sensitivität (Accruals vs. Benchmarks); Seltenheit von Haftungsfällen	(Choi et al., 2010, S. 2); (Joha, 2018, S. 148); (García-Blandón et al., 2021, S. 1393); (Günther, 2020, S. 53)

Effektrichtung (Netto)	Richtung des behaupteten Effekts auf Prüfungsqualität	+ (Qualität steigt), **0** (kein klarer Effekt), – (Qualität sinkt/ Nebenwirkungen dominieren)	(IDW, 2021, S. 5); (Völker, 2021, S. 122)
Nebenwirkungen/ Trade-offs	Welche Sekundäreffekte werden berichtet?	Marktkonzentration; Versicherbarkeit/Prämien; Angebotsverknappung; defensive auditing; Nachwuchsprobleme	(Deutsches Aktieninstitut, 2021, S. 6, 16–17); (Sethe, 2012, S. 122, 128)
Zuordnung Evidenzblock	Brücke zu 5.3–5.5	Block A (pro), Block B (contra/ Nebenwirkungen), Block C (Interdependenzen/ gemischt)	(Deutsches Aktieninstitut, 2021, S. 16); (García-Blandón et al., 2021, S. 1378, 1384)

Anhang B: Literaturmatrix

Quelle	Evidenzart	Haftungs-/Regulierungsvariable	Qualitätsmaß/Proxy	Mechanismus (Kanal)	Proxy-/Messhinweis	Effektrichtung	Kontext/Moderatoren	Block
Pfaff & Ruffing-Straube (2021, S. 240–241)	konzeptionell/übersichtsartig	Litigation Risk (Klage-/Haftungsrisiko)	indirekt (Diskussion: Effort vs. Fees)	Effort reagiert auf Litigation- und Reputationsrisiko; Reaktion kann auch über Fees laufen	Hinweis: Anpassung eher über Honorare als über Effort	gemischt	allgemeine Wirklogik, kein einzelnes Land im Fokus	A/B
Choi et al. (2010, S. 2–3, 7–8)	empirisch	ökonomische Bindung (quasi-rents/abnormal fees) als Anreizproblem (implizit: Unabhängigkeit)	Audit Fees (insb. abnormal fees) als Proxy	Fees als Preis- und Abhängigkeitskanal; Abhängigkeit kann Qualität mindern	Fees sind Mischgröße (Effort + Litigation Risk + Normal Profit)	– (bei abnormal fees)	Mandantenbeziehung/ Unabhängigkeit im Fokus	B (Mechanismus)
Joha (2018, S. 148)	konzeptionell/empirie-nah	Haftungs-/Reputationsrisiko als Bestandteil der Honorarlogik	Audit Fees	Fees enthalten (auch) Risikoprämien → erschwerte Qualitätsableitung	Risikoprämien erschweren Rückschluss auf Qualitätsdifferenzen	0 (methodischer Hinweis)	Kontextabhängig (je nach Haftungsumfeld)	C (Mess-/Proxykritik)

García-Blandón et al. (2021, S. 1378, 1380–1381, 1384, 1393)	empirisch	Joint Audit (Regulierungs-/Governanceinstrument)	Accruals und Earnings Benchmarks	„Four eyes“ vs. Free-Riding	Proxy-Sensitivität: Accruals vs. Benchmarks liefern teils andere Signale; Literatur misst stark über Accruals	0 bis – (proxyabhängig)	institutionelle Unterschiede als möglicher Einflussfaktor	C
Downar et al. (2021, S. 3)	empirisch/vergleichend argumentierend	begrenzte Haftung (cap) vs. Reputationsrisiko	indirekt (Reputation als Sanktionskanal)	Reputation kann monetäre Haftung teilweise substituieren	Fokus auf Kontext: Haftung begrenzt, Reputation dennoch stark	gemischt	Substitution abhängig von Reputationsmechanismen	C
Völker (2021, S. 108, 122)	empirisch	Haftungsrisiko/Haftungsintensität (über institutionelle Unterschiede gedacht)	je nach Studie (Proxy-Diskussion in Arbeit; hier: Qualitätsniveau/Partner)	stärkere Rechtsinstitutionen erhöhen Haftungs- und Reputationsdruck → höhere Qualität	Kontextabhängigkeit zentral	+ (unter starken Institutionen)	Stärke der Rechtsinstitutionen als Moderator	A
Günther (2020, S. 1, 53, 77)	empirisch & konzeptionell	Haftung/Haftungsfälle; Haftung als Präventivfunktion	GCM; Haftungsfälle (selten, aber präzise)	Haftung als Anreiz-/Präventionsmechanismus; Qualitätsmessung über Outputnähe	Haftungsfälle: hohe Genauigkeit, aber selten → eingeschränkte Nutzbarkeit	eher + (Anreizlogik)	Datengrundlage/Beobachtbarkeit limitiert	A/C

Langenbucher et al. (2020, S. 14–17)	Policy/konzeptionell	Liability caps (DE: 4 Mio.); private enforcement fehlt; Empfehlung: caps erhöhen	keine einheitliche Proxy-Messung (Reformargument)	höhere monetäre Sanktion → stärkere ex-ante-Anreize (Skepsis/Unabhängigkeit)	verweist auf Durchsetzungsproblem (private enforcement)	+ (Policyposition)	Durchsetzung/Litigation-Kultur als Bedingung	A
Martani et al. (2021, S. 3–5, 14)	empirisch	Rotation (audit firm/partner rotation)	Audit Quality (studienintern operationalisiert)	Rotation als Unabhängigkeits-/Qualitätshebel; heterogene Wirkung	Aufsicht/Enforcement als Moderator (niedrigere Aufsicht verändert Kosten/Risiken)	+ (heterogen)	Big-4 vs. Non-Big-4; Aufsichtsniveau	A/C
Maca & Tirdea (2025, S. 246, 249)	empirisch/marktbezogen	EU-Auditreform: Rotation, NAS-Restriktionen, Fee Caps	Investorenreaktion (Wahrnehmungsproxy); Evidenzlage zu Rotation	Bündelregulierung; Wirkung nicht isoliert	Evidenz zur Rotation als „scant“ beschrieben	gemischt	EU-Kontext; Markt-/Reformbündel	C
Deutsches Aktieninstitut (2021, S. 6, 16–17)	policy-/praxisnah	Haftung + Rotation + NAS-Verbot (Bündel)	keine einheitliche Proxy-Messung (Argumentation über Markt- und Qualitätsfolgen)	Interdependenzen treiben Konzentration; Qualität wird nicht verbessert	Isolierte Kausalattribution schwierig (Bündel)	– / 0	Reformbündel als Moderator; Marktaustritt/Nachwuchs	B/C

IDW (2021, S. 1, 5, 15–16)	Positionspapier	Haftungsverschärfung (u. Versicherbarkeit/Prämien)	keine einheitliche Proxy-Messung (normativ)	Nebenwirkungen: Konzentration, Versicherbarkeit, Angebotsverknappung, Nachwuchs	keine Proxy-Diskussion, eher Wirkungsbehauptungen	–	Berufsstands-/Marktstrukturperspektive	B
Sethe (2012, S. 122, 128)	konzeptionell/rechtspolitisch	Haftungsrecht als Eintrittsbarriere; zu hohe Haftung → Fehlanreize	keine einheitliche Proxy-Messung	Konzentration/Versicherbarkeit; defensive auditing statt Prüfungsaufgabe	betont Nebenwirkungskanal, nicht Messung	–	Marktstruktur und Versicherbarkeit	B
Haucap et al. (2022, S. 30)	policy-/gutachterlich	Haftungssummen als Markteintrittsbarriere (PIE)	keine einheitliche Proxy-Messung	Wettbewerbs-/Konzentrationseffekte	keine Proxy-Diskussion	– / gemischt	Segment PIE, Anbieterstruktur	B
Deutscher Bundestag, Finanzausschuss (2021, S. 7, 13, 16)	Parlamentsmaterial	Haftungserhöhung; Versicherbarkeit; flexible/proportionale Grenzen	keine einheitliche Proxy-Messung	Anreizargument (Sorgfalt) vs. Versicherbarkeit; Designoptionen (fix vs. flexibel)	keine Proxy-Diskussion	gemischt	Reformdesign als Moderator	A/C