Die Umsetzung europäischer Nachhaltigkeitsregulierung

Empirische Beiträge zur Prüfung doppelter Wesentlichkeit,
Vergütungsberichterstattung und EU-Taxonomie

DISSERTATION

zur Erlangung des akademischen Grades

"Doctor rerum politicarum" (Dr. rer. pol.)

vorgelegt von

Jonas Maiwald

Technische Universität Dortmund

Fakultät Wirtschaftswissenschaften

Professur für Internationale Rechnungslegung und Wirtschaftsprüfung

Dortmund

2026

Inhaltsverzeichnis

Inhaltsverzeichnis II

Tabellenverzeichnis VI

Abbildungsverzeichnis VII

Abkürzungsverzeichnis VIII

Danksagung XIII

1. Einleitung 1

1.1 Problemstellung und Zielsetzung 1

1.2 Konzeptioneller Rahmen und Theorie 7

1.3 Forschungsbeiträge 15

1.3.1 Die Prüfung der doppelten Wesentlichkeitsanalyse im Kontext der Nachhaltigkeitsberichterstattung 17

1.3.2 Vergütungsberichterstattung der HDAX-Unternehmen nach ARUG II: Formale Aspekte und textanalytische Erkenntnisse 24

1.3.3 Executive Board Gender Diversity and ESG Performance: Evidence from EU Taxonomy Disclosures 28

1.4 Gemeinsamer Beitrag 32

Literaturverzeichnis 35

2. Die Prüfung der doppelten Wesentlichkeitsanalyse im Kontext der Nachhaltigkeitsberichterstattung 59

2.1 Publikationsdetails 59

2.2 Einführung 60

2.3 Entwicklung der doppelten Wesentlichkeitsanalyse 62

2.3.1 Zunehmende Bedeutung und Regulierung der Nachhaltigkeitsberichterstattung 62

2.3.2 Funktionsweise und Umsetzung der doppelten Wesentlichkeitsanalyse gem. CSRD bzw. ESRS 64

2.3.3 Forschungsstand zur Wesentlichkeit im Kontext der Nachhaltigkeitsberichterstattung65
2.4 Prüfung der doppelten Wesentlichkeitsanalyse im Kontext wissenschaftlicher Theorien ...66
2.4.1 Externe Prüfungspflicht von Nachhaltigkeitsberichten und damit einhergehende Herausforderungen....66
2.4.2 Theoretische Begründung der externen Prüfungspflicht 68
2.5 Empirisches Forschungsdesign 70
2.5.1 Methode der qualitativen Experteninterviews 70
2.5.2 Auswahl der Interviewpartner*innen und Interviewdurchführung....71
2.5.3 Datenauswertung....72
2.6 Ergebnisse der empirischen Untersuchung 75
2.6.1 Notwendigkeit der Prüfung & Status quo....75
2.6.2 Erfahrungswerte der Prüfer und einschlägige Standards 77
2.6.3 Risiken 78
2.6.4 Prüfungsplanung 80
2.6.5 Prüfungsdurchführung 82
2.6.6 Prüfungshandlungen 84
2.6.7 Prüfungssicherheit....87
2.6.8 Bestätigungsvermerk....88
2.7 Schlussbetrachtung....89
2.7.1 Diskussion der Ergebnisse 89
2.7.2 Limitation und Implikation für Forschung und Praxis 92
Summary 93
Literaturverzeichnis 95

3. Vergütungsberichterstattung der HDAX-Unternehmen nach ARUG II: Formale Aspekte und textanalytische Erkenntnisse....105

3.1 Publikationsdetails 105
3.2 Einleitung 106

3.3 Hintergrund 106
3.3.1 Gesetz zur Umsetzung der zweiten Aktionärsrechterichtlinie (ARUG II) 106
3.3.2 Deskriptive Studien zur Vorstandsvergütung im deutschsprachigen Raum 110
3.3.3 Textanalytische Studien in der Rechnungslegungsforschung und theoretische Fundierung 111
3.4 Methodik 115
3.4.1 Untersuchungsgegenstand und Sample 115
3.4.2 Datenerhebung und Aufbereitung 116
3.5 Ergebnisse der Untersuchung und Diskussion 119
3.5.1 Ort der Veröffentlichung 119
3.5.2 Umfang, Abbildungen und weitere formale Gestaltung 120
3.5.3 Prüfung der Berichte 2021 und 2022 123
3.5.4 Lesbarkeit 124
3.6 Limitation 130
3.7 Schlussbetrachtung 130
Literaturverzeichnis 132

4. Executive Board Gender Diversity and ESG Performance: Evidence from EU Taxonomy Disclosures 140

4.1 Publication Details 140
4.2 Introduction 141
4.3 Theoretical Foundation and Hypotheses 146
4.3.1 Theoretical Framework 146
4.3.2 Literature and Hypothesis Development 149
4.4 Research Design 154
4.4.1 Sample Selection 154
4.4.2 Variable Description 156
4.4.3 Model Specification 158

4.5 Results 162
4.5.1 Descriptive Statistics 162
4.5.2 Random Effects Model Results 163
4.6 Additional Analyses and Robustness Checks 167
4.6.1 Additional Analysis: Interaction with CO_2-Intensive Companies 167
4.6.2 Robustness Check: Fractional Logit Model 170
4.7 Discussion 174
4.8 Conclusion 178
Appendix 180
References 183

Tabellenverzeichnis

Tabelle 2.1: Angaben zu den Interviewpartnern 72
Tabelle 3.1: Anzahl der Wörter in den Vergütungsberichten 121
Tabelle 3.2: Die Lesbarkeit der englischsprachigen Vergütungsberichte 125
Tabelle 3.3:Die Lesbarkeit der deutschsprachigen Vergütungsberichte 128
Table 4.1: Sample composition by industry 156
Table 4.2: Descriptive Statistics for Continuous Variables 162
Table 4.3: Descriptive Statistics of the Independent Variables (Dummies) 163
Table 4.4: Random Effects Model 165
Table 4.5: Random Effects Model: Main Results vs. CO2 Intensity 168
Table 4.6: Fractional Logit Model 172
Table 4.7: Variable Definitions, Measurements and Descriptions 180

Abbildungsverzeichnis

Abbildung 1.1: Konzeptioneller Rahmen 16
Abbildung 2.1: Datenauswertung (In Anlehnung an Kuckartz/Rädiker 2024) 73
Abbildung 2.2:Ausschnitt des Kategoriensystems 74
Abbildung 3.1: Branchenverteilung der Unternehmen 116
Abbildung 3.2: Umfang und Gestaltung der Berichte 123
Abbildung 3.3: Entwicklung der Mittelwerte der Lesbarkeitsindizes FRE und FOG 126
Abbildung 3.4: Entwicklung der Mittelwerte der Lesbarkeitsindizes LIX und FRE_{Amstad} 129

Abkürzungsverzeichnis

ABl.	Amtsblatt der Europäischen Union
Abschn.	Abschnitt
Abs.	Absatz
AG	Aktiengesellschaft
AICPA	American Institute of Certified Public Accountants
AktG	Aktiengesetz
Art.	Artikel
ARUG II	Gesetz zur Umsetzung der zweiten Aktionärsrechterichtlinie
Aufl.	Auflage
BaFin	Bundesanstalt für Finanzdienstleistungsaufsicht
BFuP	Betriebswirtschaftliche Forschung und Praxis
BGD	Board Gender Diversity
bspw.	beispielsweise
bzw.	beziehungsweise
CapEx	Capital Expenditure
CD	Cross-Sectional Dependence
CDP	Carbon Disclosure Project
CD&A	Compensation Discussion and Analysis
CEO	Chief Executive Officer
CIMA	Chartered Institute of Management Accountants
CMM	Critical Mass Model
CO_2	Kohlenstoffdioxid
CSR	Corporate Social Responsibility
CSRD	Corporate Sustainability Reporting Directive
CSR-RUG	CSR-Richtlinie-Umsetzungsgesetz

DAX	Deutscher Aktienindex
DCGK	Deutscher Corporate Governance Kodex
DRSC	Deutsches Rechnungslegungs Standards Committee
EFRAG	European Financial Reporting Advisory Group
EG	Europäische Gemeinschaft
e.g.	exempli gratia
ErwG	Erwägungsgrund
ESG	Environmental, Social and Governance
ESRS	European Sustainability Reporting Standards
et al.	et alia
EU	Europäische Union/ European Union
EURO STOXX	Aktienindex der Eurozone
EY	Ernst & Young
f.	folgende
ff.	fortfolgende
FOG	Gunning Fog Index
FRE	Flesch Reading Ease
FREAmstad	Flesch Reading Ease in der Adaption nach Amstad
FTSE	Financial Times Stock Exchange
GAAP	Generally Accepted Accounting Principles
GHG	Greenhouse Gas
GRI	Global Reporting Initiative
GVIF	Generalized Variance Inflation Factor
HDAX	Hundert Deutscher Aktienindex
HGB	Handelsgesetzbuch
IAASB	International Auditing and Assurance Standards Board

ID	Identificator
IDW	Institut der Wirtschaftsprüfer
IFAC	International Federation of Accountants
IFRS	International Financial Reporting Standards
IPCC	Intergovernmental Panel on Climate Change
IROs	Impacts, Risks and Opportunities
IRZ	Zeitschrift für Internationale Rechnungslegung
ISAE	International Standard on Assurance Engagements
ISSA	International Standard on Sustainability Assurance
ISSB	International Sustainability Standards Board
KAM(s)	Key Audit Matter(s)
KGaA	Kommanditgesellschaft auf Aktien
KoR	Zeitschrift für internationale und kapitalmarktorientierte Rechnungslegung
KPI(s)	Key Performance Indicator(s)
LIX	Lesbarkeitsindex nach Björnsson
LSEG	London Stock Exchange Group
MAXQDA	Software zur qualitativen Datenanalyse
MDAX	Mid-Cap-DAX
MD&A	Management Discussion and Analysis
NFRD	Non-Financial Reporting Directive
No	Number
Nr.	Nummer
Obs.	Observations
OLS	Ordinary Least Squares
OpEx	Operational Expenditure
PRI	Principles for Responsible Investment

Q	quartile
QMLE	Quasi-Maximum Likelihood Estimation
RDT	Resource Dependence Theory
RL	Richtlinie
Rn.	Randnummer
ROE	Return on Equity
S&P	Standard & Poor‘s
S.	Seite
SD	Standard deviation
SDAX	Small-Cap-DAX
SDGs	Sustainable Development Goals
SE	Societas Europaea
SEC	Securities and Exchange Commission
SFDR	Sustainable Finance Disclosure Regulation
sog.	sogenannte
UK	United Kingdom
UN	United Nations
UNFCCC	United Nations Framework Convention on Climate Change
U.S.	United States
SRD II	Shareholder Rights Directive II
Tz.	Textziffer
TÜV	Technischer Überwachungsverein
Var	Variance
VCOV	Variance-Covariance-Matrix
vgl.	vergleiche
Vol.	Volume

VSME	Voluntary Standard for Small and Medium-Sized Enterprises
WBCSD	World Business Council for Sustainable Development
WCED	World Commission on Environment and Development
WPg	Die Wirtschaftsprüfung
z.B.	zum Beispiel
ZfU	Zeitschrift für Umweltpolitik & Umweltrecht

Danksagung

Die vorliegende Dissertation ist das Ergebnis meiner Forschung am Lehrstuhl Internationale Rechnungslegung und Wirtschaftsprüfung der Technischen Universität Dortmund. Allen, die mich auf diesem Weg fachlich wie persönlich begleitet haben, gilt mein aufrichtiger Dank.

An erster Stelle danke ich meiner Doktormutter, Frau Prof. Dr. Christiane Pott, für ihr Vertrauen, die gewährten Freiräume und ihre kontinuierliche Unterstützung während meiner gesamten Zeit am Lehrstuhl. Ihre empathische und aufgeschlossene Art hat ein Arbeitsumfeld geschaffen, das eigenständiges Denken und persönliche Entwicklung gefördert hat und in dem ich fachlich wie persönlich wachsen konnte. Mein Dank gilt ebenso Herrn Prof. Dr. Max Pflitsch für die Übernahme des Zweitgutachtens sowie Frau Prof. Dr. Tessa Flatten für ihre Mitwirkung in der Prüfungskommission. Danken möchte ich außerdem Baker Tilly und insbesondere Martin für die finanzielle Förderung meiner Stelle, die es mir ermöglicht hat, wertvolle Praxiserfahrung mit wissenschaftlicher Forschung zu verbinden.

Herzlich danken möchte ich dem gesamten Lehrstuhlteam für die schöne gemeinsame Zeit und die vielen bereichernden Gespräche. Besonders danke ich meinen Koautoren Fabian und Björn, mit denen ich zahlreiche lange Abende damit verbracht habe, Analysen zu diskutieren, Ideen zu schärfen und aus ersten Gedanken Forschungsprojekte zu entwickeln. Diese intensive und zugleich vergnügliche Zusammenarbeit hat den wissenschaftlichen Weg ebenso produktiv wie unterhaltsam gemacht.

Von Herzen danke ich meiner Familie für den bedingungslosen Rückhalt auf diesem Weg. Mein besonderer Dank gilt meiner Frau Christina, mit der ich nicht nur die Herausforderungen des akademischen Alltags geteilt, sondern auch einige der wichtigsten Meilensteine unseres Lebens gefeiert habe, darunter unsere Hochzeit und die Geburt unserer Tochter Chloe. Deine Gelassenheit und dein Zuspruch haben mir in schwierigen Phasen Kraft gegeben und mich daran erinnert, worauf es wirklich ankommt. Mein tiefster Dank gilt außerdem meinen Eltern, deren bedingungslose Liebe und beständige Unterstützung mich mein ganzes Leben lang begleitet haben. Besonders dankbar bin ich für das entspannte und druckfreie Umfeld, das ihr mir stets geboten habt. Danken möchte ich auch meinen drei Brüdern, die stets für die nötige Ablenkung gesorgt haben, nicht selten auf der Baustelle. Ebenso herzlich danke ich meinen Schwiegereltern, die mich von Beginn an selbstverständlich aufgenommen und diesen Weg mit großem Interesse und praktischer Unterstützung begleitet haben.

Euer Vertrauen, eure klugen Ratschläge und eure Zuversicht in mich waren ein wesentliches Fundament für den Abschluss dieser Arbeit.

1. Einleitung

1.1 Problemstellung und Zielsetzung

Die Begrenzung des Klimawandels ist politisch als vorrangige Aufgabe gesetzt (United Nations, 2015; Europäische Kommission, 2019). Empirisch begründet wird diese Priorisierung durch den Zustand des Erdsystems, dessen Belastungsgrenzen in mehreren Dimensionen bereits überschritten sind (Richardson et al., 2023; Rockström et al., 2023). Der IPCC (2023) bestätigt den menschengemachten Temperaturanstieg und zeigt, dass Risiken und Auswirkungen mit jedem Erwärmungsszenario zunehmen. Die Intensivierung von Extremereignissen ist bereits nachweisbar (Min et al., 2011) und setzt sich mit weiterer Erwärmung fort (Fischer & Knutti, 2015). Bereits im Bereich von 1,5 bis 2 °C globaler Erwärmung steigt das Risiko, klimatische Kipppunkte zu überschreiten, also Schwellen, ab denen geringe weitere Erwärmung selbstverstärkende und häufig irreversible Zustandsänderungen auslöst (Lenton et al., 2008; Lenton et al., 2019; Armstrong McKay et al., 2022; Wunderling et al., 2024). Unter der Zielsetzung des Pariser Abkommens und in integrierten Bewertungsmodellen, die solche Kipppunkte berücksichtigen, erweist sich eine frühzeitige und schnelle Reduktion der CO_2-Emissionen als optimale Strategie (Cai et al., 2016).

Ökonomisch drohen langfristig erhebliche Wohlfahrts- und Bruttoinlandsproduktverluste, die mit der Erwärmung nichtlinear ansteigen (Burke et al., 2015; Hsiang et al., 2017). Entscheidend ist dabei weniger die abstrakte Kosten-Nutzen-Abwägung als die Beobachtung, dass CO_2 aktuell systematisch unterbepreist ist. Rennert et al. (2022) revidieren die sozialen Kosten von CO_2 auf 185 US-Dollar pro Tonne und damit auf das 3,6-Fache des bisherigen US-Referenzwerts. Tol (2023) bestätigt diesen Trend in einer Metaanalyse publizierter Schätzungen und kommt zu einem unmissverständlichen Schluss. Die tatsächlich erhobenen CO_2-Preise lägen nahezu überall unterhalb des geschätzten sozialen Kostenwerts, weshalb eine deutliche Anhebung geboten sei. Diese Feststellung ist für die vorliegende Dissertation von zentraler Bedeutung. Internalisiert CO_2-Bepreisung die externen Kosten nicht hinreichend, gewinnen alternative Steuerungsmechanismen wie die Lenkung von Kapital über regulatorische Informationspflichten an Gewicht.

Für Unternehmen und Investoren macht dies Klimainformationen unmittelbar entscheidungsrelevant, und zwar über mindestens vier Risikokanäle. Zu nennen sind physische Risiken in Form von Produktions- und Lieferkettenstörungen durch Extremwetter, Übergangsrisiken durch Politik- und Technologiewandel, Haftungsrisiken sowie Reputationsrisiken gegenüber Stakeholdern. Campiglio et al. (2023) systematisieren in ihrem Review diese Übertragungskanäle theoretisch und tragen die empirische Evidenz dafür zusammen, dass sowohl physische als auch Übergangsrisiken zu

Neubewertungen von Vermögenswerten führen. Für den europäischen Kapitalmarkt spezifisch dokumentieren Bua et al. (2024) signifikante Risikoprämien für beide Risikotypen seit 2015. Dass diese Risiken für Investoren schwer zu quantifizieren sind, zeigen Ilhan et al. (2021) in ihrer Optionsmarktanalyse. Krueger et al. (2020) untermauern diesen Befund mit einer Befragung institutioneller Investoren, wonach regulatorische Klimarisiken zwar als bereits materialisiert wahrgenommen werden, aber nicht vollständig in die Bewertungen einfließen.

Globalpolitisch ist der Handlungsrahmen durch das Pariser Abkommen vorgegeben (UNFCCC, 2015). Vor diesem Hintergrund rücken Kapitallenkung und Informationsinfrastruktur in den Mittelpunkt. Damit Märkte Kapital systematisch in transformationswirksame Aktivitäten lenken können, benötigen Investoren belastbare Informationen über unternehmensspezifische Auswirkungen, Risiken und Umsetzungsfortschritt (Battiston et al., 2021; Steuer & Tröger, 2022). Ob Offenlegung diese Lenkungswirkung tatsächlich entfaltet, ist jedoch fraglich. Ameli et al. (2020) zeigen für institutionelle Investoren, dass die den Transparenzinitiativen zugrunde liegende Annahme, offengelegte Klimarisiken führten zu entsprechend angepasstem Verhalten des Finanzsektors, weder theoretisch noch empirisch trägt. Kölbel et al. (2020) kommen in ihrem Review zu dem Ergebnis, dass die Wirksamkeit des Kapitalallokationsmechanismus nur teilweise belegt ist. Offenlegung allein greift damit zu kurz. Ihre Wirkung hängt vielmehr davon ab, ob die bereitgestellten Informationen innerhalb der Unternehmen in Entscheidungen überführt werden, womit Corporate Governance als Hebel der Umsetzung in den Blick gerät (Aguilera et al., 2021; Ludwig & Sassen, 2022).

Die Bereitstellung belastbarer Informationen setzt an den Steuerungs- und Kontrollstrukturen der Unternehmen an. Corporate Governance bezeichnet in der finanzwirtschaftlichen Grundlagenliteratur die Gesamtheit der Mechanismen, mit denen Kapitalgeber sicherstellen, dass die überlassenen Mittel im Sinne ihrer Interessen eingesetzt werden und ein Mittelrückfluss erfolgt (Shleifer & Vishny, 1997). Ausgangspunkt ist die aus der Trennung von Eigentum und Kontrolle resultierende Prinzipal-Agenten-Beziehung, in der Informationsasymmetrien und divergierende Interessenlagen opportunistisches Verhalten der Unternehmensleitung ermöglichen (Jensen & Meckling, 1976; Shleifer & Vishny, 1997). Als zentrale Mechanismen gelten der rechtliche Anlegerschutz, die Eigentümerstruktur, die Überwachung durch das Aufsichtsorgan, anreizkompatible Vergütungssysteme sowie die Offenlegung entscheidungsrelevanter Informationen (Shleifer & Vishny, 1997). Diese kapitalgeberzentrierte Perspektive greift für Nachhaltigkeitsfragen jedoch zu kurz, da die Auswirkungen unternehmerischen Handelns auch Anspruchsgruppen außerhalb des Kapitalgeberkreises betreffen (Freeman, 1984, S. 46). Die international maßgeblichen OECD-Grundsätze der Corporate Governance tragen dieser Erweiterung Rechnung, indem sie Nachhaltigkeit und Resilienz seit der Fassung 2023 als

eigenständigen Grundsatz führen und den Governance-Rahmen darauf verpflichten, Anreize für nachhaltigkeitsbezogene Entscheidungen und ein entsprechendes Risikomanagement zu setzen sowie wesentliche Nachhaltigkeitsinformationen offenzulegen (OECD, 2023).

Inhaltlich wird das Zielsystem, an dem sich diese Erwartungen ausrichten, über drei Begriffstraditionen aufgespannt. Corporate Social Responsibility (CSR) umfasst in der klassischen Konzeption die Gesamtheit ökonomischer, rechtlicher, ethischer und philanthropischer Verantwortlichkeiten, die eine Gesellschaft an Unternehmen heranträgt (Carroll, 1979; Carroll, 1991). Die Triple Bottom Line überträgt diesen Gedanken in eine Messlogik, indem sie die ökonomische Ergebnisgröße um eine ökologische und eine soziale Dimension ergänzt und den Unternehmenserfolg dreidimensional bestimmt (Elkington, 1998). Das Konzept der Environmental, Social and Governance (ESG)-Kriterien ist demgegenüber kapitalmarktnah entstanden und ordnet Nachhaltigkeitsaspekte den Kategorien Umwelt, Soziales und Unternehmensführung zu, um sie für Investitions- und Bewertungsentscheidungen operationalisierbar zu machen. Die Abgrenzung der Begriffe bleibt in der Literatur unscharf und werden häufig synonym verwendet (Gillan et al., 2021; Sellhorn & Wagner, 2022). Für die vorliegende Arbeit ist die Unterscheidung gleichwohl bedeutsam. Während CSR und Triple Bottom Line primär normative Verantwortungszuschreibungen formulieren, adressiert ESG die Messung und Vergleichbarkeit nachhaltigkeitsbezogener Merkmale aus Adressatenperspektive.

Immer mehr Unternehmen adressieren den Informationsbedarf zunächst freiwillig mithilfe von ESG-Berichterstattung, lange bevor eine Berichtspflicht bestand (Kolk, 2008; Simnett et al., 2009; Hahn & Kühnen, 2013; KPMG, 2022). Theoretisch lässt sich dieses freiwillige Verhalten durch ein Zusammenspiel aus Stakeholder-, Legitimitäts-, Signaling-Theorie erklären. Nach Suchman (1995) lassen sich dabei drei Legitimitätsformen unterscheiden pragmatische, moralische und kognitive Legitimität, die jeweils unterschiedliche Kommunikationsstrategien nahelegen. Unternehmen adressieren so die Erwartungen salienzstarker Anspruchsgruppen, senden Qualitätssignale in den Kapitalmarkt und reduzieren Informationsasymmetrien. Hahn und Kühnen (2013) identifizieren in ihrer Analyse von 178 Studien entsprechend Legitimitäts-, Stakeholder-, Signaling- und Institutionentheorie als die vier zentralen theoretischen Anker der Nachhaltigkeitsberichterstattung. Rouf und Siddique (2023) dokumentieren anhand von 55 Studien (2017–2021), dass Legitimitäts-, Agency- und Stakeholder-Theorie die dominierenden Erklärungsansätze für freiwillige Unternehmensberichterstattung sind, während neo-institutionalistische, Signaling-, Resource-Dependence- und Impression-Management-Theorien bislang unterrepräsentiert sind. Zamil et al. (2023) bestätigen diesen Befund mit einer breiteren Analyse von135 Studien.

Aus der Informationsökonomik lässt sich zudem eine normative Erwartung ableiten. Demnach senkt eine höhere Offenlegungsqualität das Informationsrisiko und die Kapitalkosten von Unternehmen, während eine verbesserte Vergleichbarkeit die Entscheidungsfindung der Informationsadressaten unterstützt (Lambert et al., 2007; De Franco et al., 2011). Christensen et al. (2022) zeigen jedoch, dass diese Erwartung für freiwillige ESG-Ratings nicht trägt, da mehr Offenlegung die Uneinigkeit zwischen Rating-Agenturen tendenziell erhöht statt reduziert. Berg et al. (2022) dokumentieren so beispielsweise eine ausgeprägte Divergenz von ESG-Ratings verschiedener Anbieter, die auf unterschiedliche Messansätze und Gewichtungen zurückzuführen ist.

Eine zentrale Herausforderung der ESG-Berichterstattung besteht demgemäß weiterhin in der geringen Vergleichbarkeit der verfügbaren Informationen (Korca et al., 2023), sowohl hinsichtlich der Quantität der offengelegten Daten als auch hinsichtlich ihrer Qualität und methodischen Konsistenz (Boiral & Henri, 2017; Ottenstein et al., 2022). Freiwillige Rahmenwerke haben diese Defizite nach den Befunden von Kim et al. (2026) nicht behoben. Die Übernahme von freiwilligen Rahmenwerken geht zwar mit einzelnen Verbesserungen einher, aber kein Rahmenwerk verbessert die Berichte über alle untersuchten Dimensionen hinweg. Die Autoren deuten dies als Folge fehlender verbindlicher Maßstäbe für die Substanz des Berichteten. Genau an dieser Lücke setzt die europäische Regulierung an. Die EU hat dies bereits mit der 2014 verabschiedeten Richtlinie 2014/95/EU (Non-Financial Reporting Directive, NFRD) aufgegriffen und erstmals eine entsprechende Berichtspflicht für bestimmte große Unternehmen von öffentlichem Interesse eingeführt.

Dass verpflichtende Berichterstattung diese Lücke schließen kann, deuten Christensen et al. (2021) in ihrer Synthese an, wobei vier Argumente angeführt werden, die eine Pflicht gegenüber freiwilliger Offenlegung rechtfertigen. Nachhaltigkeitsinformationen erzeugen positive Externalitäten, die einzelne Unternehmen bei freiwilliger Offenlegung nicht internalisieren, verbindliche Standards erhöhen Vergleichbarkeit und senken die Verarbeitungskosten der Adressaten, eine Pflicht wirkt als Selbstbindung gegen opportunistisches Schweigen in schlechten Jahren, und sie erlaubt es, Berichterstattung an Prüfung und Durchsetzung zu koppeln. Zugleich betonen die Autoren, dass die Kosten der Pflicht und ihre Konsequenzen für das Unternehmensverhalten von der konkreten Ausgestaltung abhängen, insbesondere von der Wahl des Wesentlichkeitskonzepts und der Intensität der Durchsetzung. Die empirische Evidenz zu bestehenden Pflichtregimen stützt diese Einschätzung. Krueger et al. (2024) zeigen für Offenlegungspflichten in 29 Ländern, dass mit der Einführung die Verfügbarkeit und Qualität von ESG-Informationen steigt, negative ESG-Vorfälle abnehmen und die Liquidität der betroffenen Aktien zunimmt, wobei die Effekte ausgeprägter sind, wenn staatliche Institutionen die Pflicht durchsetzen und keine Comply-or-explain-Ausnahmen bestehen. Grewal et

al. (2026) belegen für die britische Pflicht zur Treibhausgasberichterstattung ab 2013, dass diese Greenwashing in Form übermäßiger Länge, unbegründeten Optimismus und vager Zusagen verringerte, und zwar nicht nur bei den verpflichtend zu berichtenden Emissionsdaten, sondern auch bei den weiterhin freiwilligen übrigen Umweltangaben. Sie führen diese Übertragung auf gestiegene Reputations- und Regulierungsrisiken sowie darauf zurück, dass die für die Pflichtberichterstattung aufgebauten Governance-Strukturen der gesamten Berichterstattung zugutekommen. Beide Studien verweisen damit auf denselben Punkt, wonach nicht die Pflicht als solche, sondern ihre Durchsetzung und die durch sie ausgelösten organisatorischen Anpassungen die beobachteten Effekte erklären.

Studien zur Wirkung der NFRD liefern ein differenziertes Bild. Cuomo et al. (2022) zeigen anhand eines großen EU-Samples (2008–2018), dass die Richtlinie zu einem signifikanten Anstieg von CSR-Transparenz und -Performance geführt hat, mit stärkeren Effekten bei kleineren, stark von Analysten verfolgten Unternehmen. Aboud et al. (2024) ergänzen diesen Befund um die Frage, ob die Berichtspflicht auch das Auseinanderfallen von berichteter und tatsächlicher Nachhaltigkeitsleistung verringert, und dokumentieren für EU-Unternehmen einen Rückgang dieses ESG-Decoupling nach Einführung der Richtlinie, der bei Unternehmen mit zuvor besonders ausgeprägter Diskrepanz am stärksten ausfällt. Allerdings relativieren Ottenstein et al. (2022) dieses positive Bild in einem wichtigen Punkt. Während die Richtlinie Berichtsumfang und Prüfungswahrscheinlichkeit nachweislich erhöht, war sie nicht der entscheidende Faktor für die Übernahme von GRI-Standards[1], also gerade der Vergleichbarkeitsdimension, um die es der Richtlinie insbesondere ging. La Torre et al. (2018) deuten die Richtlinie als Legitimationsinstrument, das zwei Ziele verfolgt, verbesserte Vergleichbarkeit und gestärkte Rechenschaftspflicht. Ihrer Analyse zufolge wirken zentrale Akteure bei der praktischen Umsetzung diesen Zielen jedoch entgegen, statt sie zu befördern. Insgesamt bestätigt sich damit, dass trotz einer allgemeinen Verbesserung der Quantität und Glaubwürdigkeit der Berichterstattung insbesondere die Vergleichbarkeit als zentrales regulatorisches Ziel weiterhin unzureichend gewährleistet war. Diese Defizite bildeten einen wesentlichen Ausgangspunkt für die Weiterentwicklung des europäischen Nachhaltigkeitsberichtsrahmens im Zuge des European Green Deal (Europäische Kommission, 2019).

Mit der Corporate Sustainability Reporting Directive (CSRD)[2] und den darauf aufbauenden European Sustainability Reporting Standards (ESRS)[3] wurde die Nachhaltigkeitsberichterstattung grundlegend

[1] Die Global Reporting Initiative (GRI) ist eine internationale, gemeinnützige Organisation, die mit ihren GRI-Standards. das weltweit am weitesten verbreitete Rahmenwerk für die freiwillige Nachhaltigkeitsberichterstattung von Unternehmen bereitstellt (Villiers et al., 2022; Bais et al., 2024).
[2] Richtlinie (EU) 2022/2464.
[3] Die vorliegende Dissertation verwendet für die konzeptionelle Darstellung in diesem Kapitel die am 3. Juli 2026 von der Europäischen Kommission verabschiedeten, überarbeiteten ESRS (Delegierte Verordnung C(2026) 5010 final), zum

ausgeweitet und stärker vereinheitlicht. Während die CSRD den Anwendungsbereich und die Berichtspflichten erweitert, sollen die ESRS durch einheitliche europäische Berichtsstandards insbesondere die zuvor dokumentierte Heterogenität und die bestehenden Qualitätsdefizite der Berichterstattung reduzieren (Alliance for Corporate Transparency, 2019/2020; Erwägungsgründe CSRD, 2022). Ergänzend etabliert die EU-Taxonomie-Verordnung[4] gem. Art. 8 einheitliche quantitative Kennzahlen zur Nachhaltigkeitsleistung wirtschaftlicher Aktivitäten, insbesondere in Form der Anteile nachhaltigkeitsbezogener Umsätze sowie von CapEx und OpEx.

Ziel der vorliegenden Dissertation ist es, einen integrierten Einblick in die Funktionsweise und die Grenzen ausgewählter EU-Instrumente zur Förderung nachhaltigen Wirtschaftens zu geben, konkret der CSRD samt ESRS, der EU-Taxonomie-Verordnung sowie des Gesetzes zur Umsetzung der zweiten Aktionärsrechterichtlinie (ARUG II) als nationale Umsetzung der Richtlinie (EU) 2017/828. Anders als CSRD und EU-Taxonomie entstammt ARUG II nicht dem Green-Deal-Instrumentarium im engeren Sinne, sondern der allgemeinen Corporate Governance Regulierung. Seit der Verpflichtung zur Kopplung von Vergütungszielen an die langfristige Entwicklung der Gesellschaft (§ 87 Abs. 1 AktG; § 162 Abs. 1 Nr. 1 AktG) ist jedoch auch hier ein zunehmender Nachhaltigkeitsbezug erkennbar.[5] Zum einen soll untersucht werden, wie diese Instrumente in der Praxis wirken, konkret in der Prüfungspraxis von Wirtschaftsprüfern, in der Vergütungsberichterstattung börsennotierter Unternehmen sowie im Zusammenhang zwischen Vorstandszusammensetzung und Taxonomie-Umsetzung. Zum anderen soll konzeptionell eingeordnet werden, inwiefern formale Regelkonformität tatsächlich mit substanzieller unternehmerischer Transformation einhergeht. Diesen Zusammenhang beleuchten die drei Beiträge aus unterschiedlichen Blickwinkeln. Beitrag 1 untersucht die Voraussetzungen glaubwürdiger Prüfungspraxis, Beitrag 2 die empirisch dokumentierte Lücke zwischen Berichtsumfang und tatsächlicher Verständlichkeit und Beitrag 3 konkrete Governance-Bedingungen, die Implementierung begünstigen können.

Damit schließt die Arbeit an eine seit Langem bekannte, aber empirisch noch immer unzureichend bearbeitete Forschungslücke an. Bereits Naciti et al. (2021) konstatieren im Rahmen einer systematischen Literaturrecherche von 468 Arbeiten, dass bislang nur wenige empirische Untersuchungen zur tatsächlichen Wirkung von Governance-Praktiken auf

Zeitpunkt der Fertigstellung dieser Arbeit noch ohne Amtsblattnummer, da der Rechtsakt der Prüfung durch Europäisches Parlament und Rat unterliegt). Beitrag 1 bezieht sich demgegenüber auf die zum Erhebungszeitpunkt geltende ursprüngliche Fassung nach Delegierter Verordnung (EU) 2023/2772, da die Interviews vor der Verabschiedung der überarbeiteten Standards durchgeführt wurden.

[4] Verordnung (EU) 2020/852.

[5] Auch die OECD-Grundsätze der Corporate Governance stellen diesen Zusammenhang her, indem sie die Ausrichtung der Organvergütung explizit mit Nachhaltigkeitsleitung in Verbindung bringen (OECD, 2023).

Nachhaltigkeitsperformance vorliegen und weiterer Forschungsbedarf besteht. Aktuellere Reviews bestätigen, dass diese Lücke fortbesteht (Aguilera et al., 2021; Buchetti et al., 2025). Die vorliegende Dissertation adressiert diese Lücke nicht durch eine einzelne Studie, sondern durch drei komplementäre empirische Zugänge, die jeweils eine andere Ebene der Umsetzung, Prüfung, Kommunikation und Implementierung in den Blick nehmen.

Das folgende Kapitel entwickelt hierfür einen konzeptionellen Rahmen, der diese drei Ebenen theoretisch verortet und ihr Zusammenspiel begründet.

1.2 Konzeptioneller Rahmen und Theorie

Die vorliegende Dissertation versteht CSRD/ESRS, EU-Taxonomie und ARUG II nicht als drei isolierte Regelwerke, sondern als Bausteine einer gemeinsamen, mehrstufigen Informationsarchitektur, die europäische Nachhaltigkeitsziele in unternehmerisches Handeln übersetzen soll. Die CSRD und ESRS definieren Inhalte, Prozesse und Prüfungsanforderungen und institutionalisieren mit der doppelten Wesentlichkeit eine zweidimensionale Relevanzbestimmung. Diese verknüpft systematisch die finanzielle Perspektive, also die Frage, wie sich Nachhaltigkeitsaspekte auf die wirtschaftliche Lage des Unternehmens auswirken, mit der Impact-Perspektive, also der Frage, wie sich unternehmerisches Handeln auf Umwelt und Gesellschaft auswirkt (ESRS 1.35 ff.; CSRD Erwägungsgrund 29). Die EU-Taxonomie stellt dazu outcome-nahe Messgrößen bereit, mit denen Transformation und Investitionsallokation auf Unternehmensebene nachvollziehbar werden (Pacces, 2021; Nipper et al., 2025). Das ARUG II verpflichtet Vorstand und Aufsichtsrat zu einem klaren und verständlichen Vergütungsbericht (§ 162 AktG) und schreibt vor, dass die Vergütungsstruktur auf eine nachhaltige und langfristige Unternehmensentwicklung auszurichten ist (§ 87 Abs. 1 AktG). Diese ursprünglich klassisch-aktienrechtlich, im Sinne der Vermeidung kurzfristiger Anreize, zu verstehende Vorgabe wird in der Unternehmenspraxis zunehmend ESG-spezifisch ausgelegt (Cohen et al., 2023; Dell'Erba & Ferrarini, 2024; Spallek & Bastini, 2026). In dieser Architektur ist Corporate Governance nicht bloß Träger der Berichterstattung, sondern Hebel der Umsetzung. Sie entscheidet über Zielsetzungen, Ressourcenallokation (Alkaraan et al., 2023; Hristov & Searcy, 2025), interne Kontroll- und Datenprozesse (Harasheh & Provasi, 2023) sowie über die Einbettung externer Assurance in unternehmensweite Steuerungsroutinen (Ludwig & Sassen, 2022). Kapazitätsbezogen hängt die Wirksamkeit dieser Instrumente zugleich davon ab, ob und wie das Leitungsgremium verfügbare Informationen aufnimmt, gewichtet und in Entscheidungen überführt. Gremien unterliegen dabei kognitiven Beschränkungen, die ihre Aufmerksamkeit selektiv auf einzelne Aufgaben lenken und dadurch begrenzen, welche Informationen überhaupt entscheidungswirksam werden (Forbes & Milliken, 1999; Khanna et al., 2014; Schiehll et al., 2023).

Beide Dimensionen finden sich in den Beiträgen dieser Arbeit wieder. Die ersten beiden Beiträge setzen mit der externen Prüfung und der Vergütungsberichterstattung an instrumentellen Ausprägungen an, während der dritte Beitrag mit der Zusammensetzung des Leitungsgremiums eine strukturelle Voraussetzung der Informationsverarbeitung untersucht. Konzeptionell bewegt sich die Arbeit damit in einem Forschungsfeld, das nachhaltigkeitsbezogene Zielsetzungen mit klassischen Steuerungs- und Kontrollmechanismen zusammenführt, dessen Wirkungszusammenhänge jedoch bislang nur unvollständig geklärt sind (Elkington, 2006; Velte, 2021a; Velte, 2022a; Tutino et al., 2025).

Die Arbeit ist theoretisch pluralistisch angelegt und greift je nach Fragestellung und Anwendungskontext der einzelnen Beiträge auf unterschiedliche komplementäre Theorieansätze zurück. Einen übergreifenden Bezugsrahmen bildet die Stakeholder-Theorie, wonach unternehmerische Legitimität und langfristige Wertschöpfung wesentlich davon abhängen, inwieweit Unternehmen in der Lage sind, heterogene Stakeholdererwartungen zu identifizieren, zu berücksichtigen und transparent zu adressieren (Freeman, 1984). Diese Perspektive ist insbesondere für die Konzeption und Prüfung der doppelten Wesentlichkeit zentral, da diese explizit über eine rein investorenorientierte Sichtweise hinausgeht (ESRS 1.4). Dass diese Erweiterung des Adressatenkreises nicht nur normativ gesetzt ist, sondern auf tatsächliche Nachfrage trifft, zeigen Hagemeier und Müller (2026) anhand von Web-Tracking-Daten zu den ersten CSRD-Berichten. Nachhaltigkeitsangaben werden demnach ähnlich häufig abgerufen wie Finanzangaben, jedoch von einem breiteren und stärker europäisch geprägten Stakeholderkreis, während Finanzangaben ein globaleres und investorenorientiertes Publikum erreichen.

Ergänzend wird die Prinzipal-Agent-Theorie (Jensen & Meckling, 1976) herangezogen, um die Rolle von Nachhaltigkeitsberichterstattung und insbesondere von Assurance-Leistungen als Instrumente zum Abbau von Informationsasymmetrien zu analysieren (vgl. den grundlegenden Literaturüberblick von Healy und Palepu, 2001). In diesem Zusammenhang wird argumentiert, dass die Prüfung nicht nur die formale Regelkonformität sicherstellt, sondern maßgeblich zur Glaubwürdigkeit und Entscheidungsnützlichkeit der berichteten Informationen beiträgt (Ioannou & Serafeim, 2017; García-Sánchez et al., 2019). Vor diesem Hintergrund adressiert ESG-Assurance die Unsicherheit der Informationsadressaten hinsichtlich der Richtigkeit, Vollständigkeit und Angemessenheit der offengelegten Nachhaltigkeitsinformationen und damit eine zentrale Schwachstelle der nichtfinanziellen Berichterstattung (Simnett et al., 2009).

Darüber hinaus werden legitimitätstheoretische und signaling-theoretische Ansätze genutzt, um freiwillige Prüfungsentscheidungen von Vergütungsberichten sowie eine verständliche Aufbereitung

der Berichte als bewusste Kommunikations- und Vertrauenssignale gegenüber Kapitalmarktteilnehmern zu interpretieren (Hemmings et al., 2020). Der europäische Kontext mit stärker kodifizierten Stakeholder-Rechten und einer politisch breiteren Akzeptanz von Nachhaltigkeitsregulierung schafft dagegen eher Anreize, Assurance als integralen Bestandteil eines institutionalisierten Transparenzregimes zu nutzen (Kolk & Perego, 2010; Simoni et al., 2020). Assurance kann also als institutionalisierter Vertrauensmechanismus, der in Anlehnung an Stakeholder-, Legitimitäts- und Prinzipal-Agenten-Theorie Informationsasymmetrien reduziert, opportunistisches Verhalten begrenzt und die Wahrnehmung der Berichte als verlässliche Grundlage wirtschaftlicher und gesellschaftlicher Entscheidungen stärkt, verstanden werden (Cuadrado-Ballesteros et al., 2017; Quick, 2024; Del Gesso & Lodhi, 2025)

Die Institutionentheorie erklärt Unterschiede zwischen der Europäischen Union und den USA als Resultat unterschiedlicher institutioneller Umwelten, in denen ESG-Assurance eingebettet ist. DiMaggio und Powell (1983) unterscheiden drei Formen institutionellen Isomorphismus. Erzwingender Isomorphismus entsteht aus politischem Einfluss und Legitimationsdruck, mimetischer Isomorphismus resultiert aus Nachahmung unter Unsicherheit, und normativer Isomorphismus geht mit Professionalisierung einher. Während sich der vorangehende Befund zu Stakeholder-Rechten und institutionalisierten Transparenzregimen auf die freiwillige Assurance-Entscheidung einzelner Unternehmen bezieht, wirkt erzwingender Isomorphismus in der Europäischen Union inzwischen unmittelbarer. Mit NFRD, CSRD und ESRS existiert ein verpflichtendes Regime, das Berichts- und zunehmend auch Prüfungspflichten rechtlich vorschreibt, statt sie lediglich durch institutionellen Druck nahezulegen (Martínez-Ferrero & García-Sánchez, 2017; Ștefănescu, 2022). In den USA dominiert hingegen weiterhin ein stärker marktorientiertes, freiwilliges Offenlegungsregime, in dem mimetische und normative Einflüsse wichtiger sind als verbindliche gesetzliche Pflichten (Casey & Grenier, 2015). Dass diese Unterschiede nicht allein regulatorisch, sondern in den Corporate-Governance-Systemen selbst angelegt sind, zeigen Aguilera et al. (2006) im Vergleich des Vereinigten Königreichs und der USA. Die stärker auf Stakeholder und langfristige Orientierung ausgerichtete britische Governance geht demnach mit einem substanzielleren Engagement institutioneller Investoren für unternehmerische Verantwortung einher, während das stärker shareholderzentrierte US-System dieses Engagement schwächer ausprägt. Für die Analyse der Umsetzung der EU-Taxonomie wird der theoretische Rahmen um ressourcenabhängige Ansätze sowie die Critical Mass-Theorie nach Kanter (1977) erweitert. Nach der Resource-Dependence-Theorie verbessern bestimmte Governance-Merkmale wie eine erhöhte Geschlechterdiversität im Vorstand den Zugang zu relevanten Ressourcen, stärken die Berücksichtigung externer Stakeholderinteressen und fördern inklusivere

Entscheidungsprozesse (Pfeffer & Salancik, 1978). Für die Berichtsqualität ist dieser Zusammenhang empirisch assoziiert (Al-Shaer & Zaman, 2016). Ob er sich auch auf die tatsächliche Implementierung regulatorischer Nachhaltigkeitsanforderungen erstreckt, ist bislang kaum untersucht und bildet den Ausgangspunkt des dritten Beitrags.

Vor dem Hintergrund dieser theoretischen Perspektiven lässt sich der Zusammenhang zwischen Corporate Governance und Nachhaltigkeitsleistung präzisieren. Eine Durchsicht der bestehenden Literatur zeigt, dass Agency-Theorie (Jensen & Meckling, 1976) und Stakeholder-Theorie (Freeman, 1984) die beiden dominierenden Perspektiven zur Erklärung dieses Zusammenhangs darstellen (Hussain et al., 2018). Beide Theorien führen dabei zu unterschiedlichen, teils komplementären Erwartungen. Während die Agency-Perspektive Governance-Mechanismen primär als Kontrollinstrumente zur Reduktion von Interessenkonflikten zwischen Management und Kapitalgebern versteht, betont die Stakeholder-Perspektive die Notwendigkeit, Governance-Strukturen so zu gestalten, dass sie ein breiteres Spektrum an Anspruchsgruppeninteressen systematisch berücksichtigen. Naciti et al. (2021) bestätigen dieses Theoriemuster im Rahmen einer systematischen Literaturrecherche von 468 Arbeiten und identifizieren dieselben beiden Ansätze als die in der Forschung am häufigsten herangezogenen Erklärungsrahmen.

Christensen et al. (2021) zeigen in ihrer Synthese, dass die ökonomischen Konsequenzen verpflichtender Nachhaltigkeitsberichterstattung höchst heterogen ausfallen und maßgeblich von Ausgestaltung und Durchsetzung abhängen, nicht von der bloßen Existenz einer Offenlegungspflicht. Dass sich CSRD, EU-Taxonomie, Sustainable Finance Disclosure Regulation (SFDR) [6] und die Pillar-3-Offenlegungspflichten[7] der EU zunehmend wechselseitig bedingen und aufeinander aufbauen, bestätigen Hummel und Jobst (2024) in ihrem Überblick über die gesamte EU-Regulierungslandschaft. Dieser Befund stützt die Wahl eines integrierten Ebenenmodells.

Bereits die frühe empirische Literatur zeigt jedoch, dass die von der Prinzipal-Agent-Theorie erwartete Reduktion von Informationsasymmetrien durch Offenlegung nicht automatisch erfolgt. Marshall und Weetman (2002) vergleichen britische und US-amerikanische Unternehmen und zeigen, dass dieselbe regulatorische Zielsetzung zu Offenlegungspflichten für Wechselkursrisiken in beiden Ländern zu unterschiedlichen Ergebnissen führt, wobei in keinem der beiden Regime die Informationsasymmetrie

[6] Die SFDR verpflichtet Finanzmarktteilnehmer zur Offenlegung, inwiefern sie Nachhaltigkeitsaspekte in ihre Finanzprodukte und Anlageentscheidungen einbeziehen (Verordnung (EU) 2019/2088)).

[7] Die Pillar-3-Offenlegungspflichten beruhen demgegenüber auf Art. 449a der Eigenkapitalverordnung, eingeführt durch die Verordnung (EU) 2019/876 und seither mehrfach aktualisiert. Die Durchführungsverordnung (EU) 2022/2453 etwa verpflichtet Kreditinstitute zur Offenlegung ihrer Umwelt-, Sozial- und Governance-Risiken gegenüber der Aufsicht und dem Markt.

vollständig abgebaut wird. Aktuellere, global angelegte Evidenz von Krueger et al. (2024) bestätigt dieses Muster. ESG-Offenlegungspflichten erhöhen die Aktienliquidität signifikant, aber nur, wenn sie von staatlichen Institutionen mit strenger Durchsetzung statt auf Comply-or-Explain-Basis implementiert werden.

Aus dieser Beobachtung leitet sich die Arbeitshypothese der Dissertation ab, wonach regulatorische Informationspflichten eine notwendige, aber keine hinreichende Bedingung für reale unternehmerische Transformation darstellen und ihre Wirksamkeit sich unter anderem auf der Ebene der Umsetzungsqualität entscheidet. Sie dient im Folgenden als Ordnungsrahmen für die drei Beiträge, ohne dass die vorliegende Arbeit die darin angelegte Kausalbeziehung selbst prüft.

Auf der ersten Ebene, Offenlegung und Assurance, institutionalisieren CSRD und ESRS mit der doppelten Wesentlichkeit eine zweidimensionale Relevanzbestimmung, deren Belastbarkeit erst durch externe Prüfung sichergestellt wird (Beitrag 1). Länder, die eine Offenlegungspflicht einführen, verzeichnen nicht nur mehr Offenlegung, sondern auch eine signifikant erhöhte freiwillige Nachfrage nach Assurance. Die Publizitätspflicht allein scheint demnach nicht zu genügen, um die Glaubwürdigkeit der berichteten Informationen herzustellen (Ioannou & Serafeim, 2017). García-Sánchez et al. (2021) untermauern dies mit einer Untersuchung von 1.939 Unternehmen über 15 Jahre. Demzufolge reduziert Assurance Decoupling-Praktiken, also das Auseinanderfallen von formal berichteter und tatsächlich umgesetzter Nachhaltigkeitsleistung, wobei die Wirkung stark von den Merkmalen des Prüfungsanbieters abhängt. Diese Entkopplung bildet ein wiederkehrendes Motiv der vorliegenden Arbeit und wird im Abschnitt zu den Grenzen des Modells theoretisch entfaltet. Al-Shaer und Zaman (2018) liefern über die Resource-Dependence-Theorie zudem eine direkte konzeptionelle Brücke zu Beitrag 3, indem sie zeigen, dass Prüfungsausschuss-Merkmale die Assurance-Entscheidung beeinflussen.

Auf der zweiten Ebene, Kommunikationsqualität, positioniert sich Beitrag 2 mit der Untersuchung der Vergütungsberichterstattung nach ARUG II. Theoretischer Ausgangspunkt ist die Obfuscation-Hypothese, wonach das Management die sprachliche Gestaltung der Berichterstattung strategisch einsetzt, um nachteilige Informationen schwerer zugänglich zu machen (Courtis, 1998, 2004; Li, 2008). Hooghiemstra et al. (2017) zeigen anhand britischer Unternehmen, dass bei exzessiver Vorstandsvergütung ein schlechter lesbarer Vergütungsbericht mit geringerem Widerspruch bei der aktionärsseitigen Abstimmung über die Vergütung, dem sogenannten Say-on-Pay-Votum, einhergeht, was die Obfuscation-Hypothese stützt. Dieser Effekt schwächt sich mit steigendem institutionellem Anteilsbesitz ab und kann sich bei mehrheitlich institutionellem Besitz sogar umkehren. Hemmings et al. (2020) gelangen für den US-amerikanischen Kontext zu einem gegenläufigen Ergebnis. Eine

bessere Lesbarkeit der Compensation Discussion and Analysis (CD&A) schwächt dort den Zusammenhang zwischen exzessiver Vergütung und negativem Abstimmungsergebnis ab, was die Autoren über die Processing-Fluency-Heuristik nach Rennekamp (2012) erklären, wonach verständlich dargebotene Informationen als glaubwürdiger wahrgenommen werden. Beide Studien dokumentieren damit eine höhere Zustimmung, führen sie jedoch auf entgegengesetzte Mechanismen zurück. Für den deutschsprachigen Raum lässt sich die Obfuscation-Hypothese bislang allerdings nicht replizieren. Thoms et al. (2020) finden für die DAX-30-Geschäftsberichte keinen systematischen Lesbarkeitsunterschied zwischen positiv und negativ konnotierten Passagen. Stellner (2022) kommt für den Wiener Prime Market zum gleichen Ergebnis. Dies begründet, die Lesbarkeitsentwicklung im deutschen Regulierungskontext zunächst deskriptiv zu erfassen, statt angelsächsische Obfuscation-Befunde vorauszusetzen.

Zwischen Ebene 2 und Ebene 3 lässt sich zudem eine theoretische Brücke konstruieren. Nadeem (2022) zeigt empirisch, dass Board-Gender-Diversity die Lesbarkeit der Jahresberichte US-amerikanischer börsennotierter Unternehmen (10-K-Berichte) verbessert. Gabaldón et al. (2018) liefern hierfür über die Information-Processing-Theorie eine kausale Erklärung, wonach Diversität im Gremium nicht direkt wirkt, sondern über die Art, wie unterschiedliches Wissen und unterschiedliche Kompetenzen im Entscheidungsprozess genutzt werden. Kanadlı et al. (2018) bestätigen diesen Vermittlungsmechanismus für die Umfassendheit und Geschwindigkeit von Gremienentscheidungen. Zhang (2010) differenziert ihn weiter und unterscheidet zwischen dem bloßen Besitz diverser Information und ihrer tatsächlichen Nutzung im Gremium.[8] Die deskriptiven Lesbarkeitsbefunde aus Beitrag 2 und die Wirkungsanalyse aus Beitrag 3 lassen sich damit auf die eingangs eingeführte kapazitätsbezogene Dimension zurückführen, also auf die Informationsverarbeitungskapazität des Leitungsgremiums (Pirson & Turnbull, 2011; Seo, 2017). Dieser Mechanismus tritt allerdings neben die zuvor erörterte Obfuscation-Perspektive und ersetzt sie nicht. Während Obfuscation Lesbarkeit als Ergebnis strategischer Absicht deutet (Courtis, 1998, 2004; Li, 2008), erklärt die Informationsverarbeitungsperspektive sie als Ausdruck organisationaler Verarbeitungsfähigkeit (Seo, 2017; Pirson & Turnbull, 2011). Welcher der beiden Mechanismen im jeweiligen Fall dominiert, lässt sich anhand der hier vorgelegten Befunde nicht bestimmen.

Auf der dritten Ebene, Implementierung und Wirkung, untersucht Beitrag 3, ob die Zusammensetzung des Vorstands mit der tatsächlichen Umsetzung der EU-Taxonomie-Anforderungen zusammenhängt

[8] Einschränkend ist anzumerken, dass Gabaldón et al. (2018) und Zhang (2010) auf norwegischen Befragungsdaten desselben Erhebungszusammenhangs beruhen und Zattoni et al. (2023) in ihrem Review von 184 Studien eine insgesamt heterogene Evidenzlage zu Board-Diversität konstatieren.

und nicht nur mit deren Offenlegung. Theoretisch stützt sich der Beitrag auf ein Zusammenspiel aus Stakeholder-Theorie (Freeman, 1984; Michelon & Parbonetti, 2012), Resource-Dependence-Theorie (Pfeffer & Salancik, 1978; Hillman & Dalziel, 2003) und Critical Mass-Theorie (Kanter, 1977). Die Stakeholder-Perspektive begründet, weshalb divers besetzte Gremien gesellschaftliche und regulatorische Nachhaltigkeitserwartungen aufmerksamer aufgreifen, die Resource-Dependence-Theorie erklärt dies über die Ressourcenfunktion von Boards als Träger von Expertise, Netzwerken und Legitimität, und die Critical-Mass-Theorie präzisiert, dass substanzielle Wirkung erst oberhalb einer Repräsentationsschwelle eintritt. Dass Taxonomie-Kennzahlen tatsächlich handlungsleitend wirken, zeigen Chrzan und Pott (2024) in zwei experimentellen Studien mit institutionellen und privaten Investoren. Die Einbeziehung taxonomiebezogener Kennzahlen beeinflusst Investitionsurteile messbar, wobei professionelle Investoren negative Informationen deutlich stärker sanktionieren als private. Schütze und Sandbaek (2025) ordnen die Taxonomie in ihre drei regulatorischen Zielsetzungen ein, die Erhöhung von Transparenz, die Reduktion von Greenwashing und die Lenkung von Kapital in nachhaltige Aktivitäten.

Die Grenze dieses Drei-Ebenen-Modells liegt in der fehlenden empirischen Verkettung zwischen den Ebenen. Keine der drei Studien testet direkt, ob bessere Prozessqualität kausal zu besserer Wirkung führt. Die Decoupling-Literatur liefert hierfür eine theoretisch fundierte Erklärung. Meyer und Rowan (1977) begründen mit ihrem Beitrag zur Organisationssoziologie, dass formale Struktur und tatsächliche Praxis systematisch auseinanderfallen können, ohne dass dies Ausdruck von Versagen, sondern von rationalem Legitimationsstreben ist. Wijen (2014) verfeinert dies zum Means-Ends-Decoupling, wonach selbst vollständige formale Umsetzung die Erreichung des intendierten Ziels nicht garantiert. Cho et al. (2015) gehen einen Schritt weiter und argumentieren, dass Nachhaltigkeitsberichterstattung unter widersprüchlichen Stakeholdererwartungen und institutionellen Zwängen strukturell auf organisierte Heuchelei und organisationale Fassaden angewiesen ist, sodass die Lücke zwischen Darstellung und Handeln nicht als Fehlfunktion, sondern als systemimmanentes Merkmal zu verstehen ist.

Für Ebene 1 zeigen García-Meca et al. (2026) anhand von STOXX-Europe-600-Unternehmen, dass Assurance nach medial berichtetem ESG-Fehlverhalten gezielt zur Legitimitätsreparatur eingesetzt wird und die Einführung der NFRD an dieser symbolischen Verwendung nichts geändert hat. Boiral und Heras-Saizarbitoria (2020) verschärfen diesen Befund anhand einer Inhaltsanalyse von 337 geprüften Nachhaltigkeitsberichten, wonach Assurance-Statements vielfach keinen materiellen und glaubwürdigen Prüfungsprozess belegen, sondern von den eigentlichen Nachhaltigkeitsfragen abgelöst bleiben. Diese Befunde stehen in Spannung zu den zuvor referierten Ergebnissen von García-

Sánchez et al. (2021), wonach Assurance Decoupling reduziert, und verweisen darauf, dass Prüfung Glaubwürdigkeit nicht unter allen Bedingungen erhöht. Für den Regelungsbereich des ARUG II, der neben dem Vergütungsbericht auch die Vergütungsstruktur erfasst, zeigt Lu (2023) anhand der FTSE-350-Unternehmen, dass ESG-basierte Vergütung ihre intendierte Steuerungsfunktion häufig verfehlt und stattdessen als Instrument des Impression Management oder der Rentenaneignung dient, da sich die Wirkung der meisten ESG-Faktoren auf den Unternehmenswert kaum objektiv messen lässt. Bebchuk und Tallarita (2022) bestätigen für die S&P-100-Unternehmen, dass sich von außen in nahezu allen Fällen nicht beurteilen lässt, ob ESG-Kennzahlen echte Anreize setzen oder lediglich leistungsunabhängige Vergütung legitimieren. Beitrag 3 liefert für diese Diagnose eigene Evidenz. Dort sind sowohl das Vorhandensein eines CSR-Ausschusses als auch die Kopplung der Vorstandsvergütung an Nachhaltigkeitskennzahlen negativ mit der betriebsausgabenbezogenen Taxonomie-Implementierung assoziiert, letztere allerdings nur im Fractional-Logit-Modell. Ein Anreizinstrument, dessen Ausgestaltung das ARUG II über § 87 Abs. 1 AktG adressiert, geht damit nicht mit höherer Umsetzung auf Ebene 3 einher, während die personelle Zusammensetzung des Vorstands es tut. Für Ebene 3 relativieren Tilbury und Sealy (2024) die Critical Mass-Annahme anhand einer vergleichenden Fallstudie dreier Boards, wonach das Erreichen der Schwelle von drei weiblichen Mitgliedern Beteiligungsbarrieren nicht zwangsläufig beseitigt, sondern Prozessverluste fortbestehen können. Diese Widersprüche werden hier nicht aufgelöst, sondern als Beleg dafür angeführt, dass Wirksamkeit je nach Operationalisierung, institutionellem Kontext und untersuchtem Mechanismus unterschiedlich ausfällt. Über diese literaturgestützten Belege hinaus benennt Beitrag 3 selbst zwei weitere Vorbehalte gegenüber dem vermuteten Governance-Effekt, die in der Zusammenfassung des Beitrags in Kapitel 1.3.3 dargestellt sind.

Schließlich verdient die aktuelle regulatorische Dynamik Erwähnung, ohne die Kernthese zu schwächen. Velte (2026) analysiert kritisch das sog. EU-Omnibus-Vereinfachungspaket[9], das CSRD und Taxonomie-Verordnung erheblich im Anwendungsbereich reduziert, und argumentiert aus stakeholder- und legitimitätstheoretischer Perspektive, dass diese Verengung dem Green-Deal-Ziel zuwiderläuft und die Legitimität der EU-Kommission belastet. Ruohonen und Tamlander (2025) ergänzen diese Einschätzung aus Sicht von Unternehmen und Prüfern und weisen darauf hin, dass die Entscheidung, es bei begrenzter Prüfungssicherheit zu belassen, die regulatorische Steuerungswirkung

[9] Mit dem Omnibus-I-Paket verfolgt die Europäische Kommission das Ziel, die Wettbewerbsfähigkeit europäischer Unternehmen durch den Abbau administrativer Lasten zu stärken. Die zur Umsetzung erlassene Richtlinie (EU) 2026/470 hebt unter anderem die Schwellenwerte der CSRD-Berichtspflicht auf mehr als 1.000 Mitarbeitende und mehr als 450 Mio. Euro Nettoumsatz an und streicht die Ermächtigung zum Erlass von Standards für Prüfungen mit hinreichender Sicherheit. Flankierend verabschiedete die Kommission am 3. Juli 2026 überarbeitete ESRS mit deutlich reduziertem Datenpunktumfang.

zusätzlich schwächt. Eine zweistufige Praktiker- und Expertenbefragung von Stanek-Kowalczyk et al. (2026) vor und nach Ankündigung des Omnibus-Pakets bestätigt dies aus der Perspektive der Berichtsersteller. Erwartet werden zwar Vereinfachung und geringere Berichtspflichten, zugleich aber eine abnehmende Bedeutung von ESG-Expertise und nur begrenzte Verbesserungen der Berichtsqualität. Diese Beobachtung schwächt die Klammer nicht, sondern bestätigt sie in neuem regulatorischem Gewand. Auch eine im Anwendungsbereich reduzierte Architektur bleibt in ihrer Wirksamkeit von der Qualität der Umsetzung abhängig, sodass sich die zentrale Frage dieser Dissertation unter dem Omnibus-Regime mit unveränderter Dringlichkeit stellt. Die folgenden Kapitel wenden sich diesen drei Ebenen im Detail zu. Kapitel 2 behandelt die Prüfungspraxis der doppelten Wesentlichkeitsanalyse, Kapitel 3 die Vergütungsberichterstattung nach ARUG II und ihre Lesbarkeitswirkung und Kapitel 4 den Zusammenhang zwischen Vorstandszusammensetzung und Taxonomie-Implementierung.

1.3 Forschungsbeiträge

Die Dissertation kombiniert drei methodisch unterschiedliche Forschungsdesigns, die in ihrer Erkenntnislogik komplementär sind. Die drei Forschungsgegenstände befinden sich in unterschiedlichen Reifestadien empirischer Erschließung und stellen methodisch unterschiedliche Anforderungen (Vgl. Abbildung 1.1).

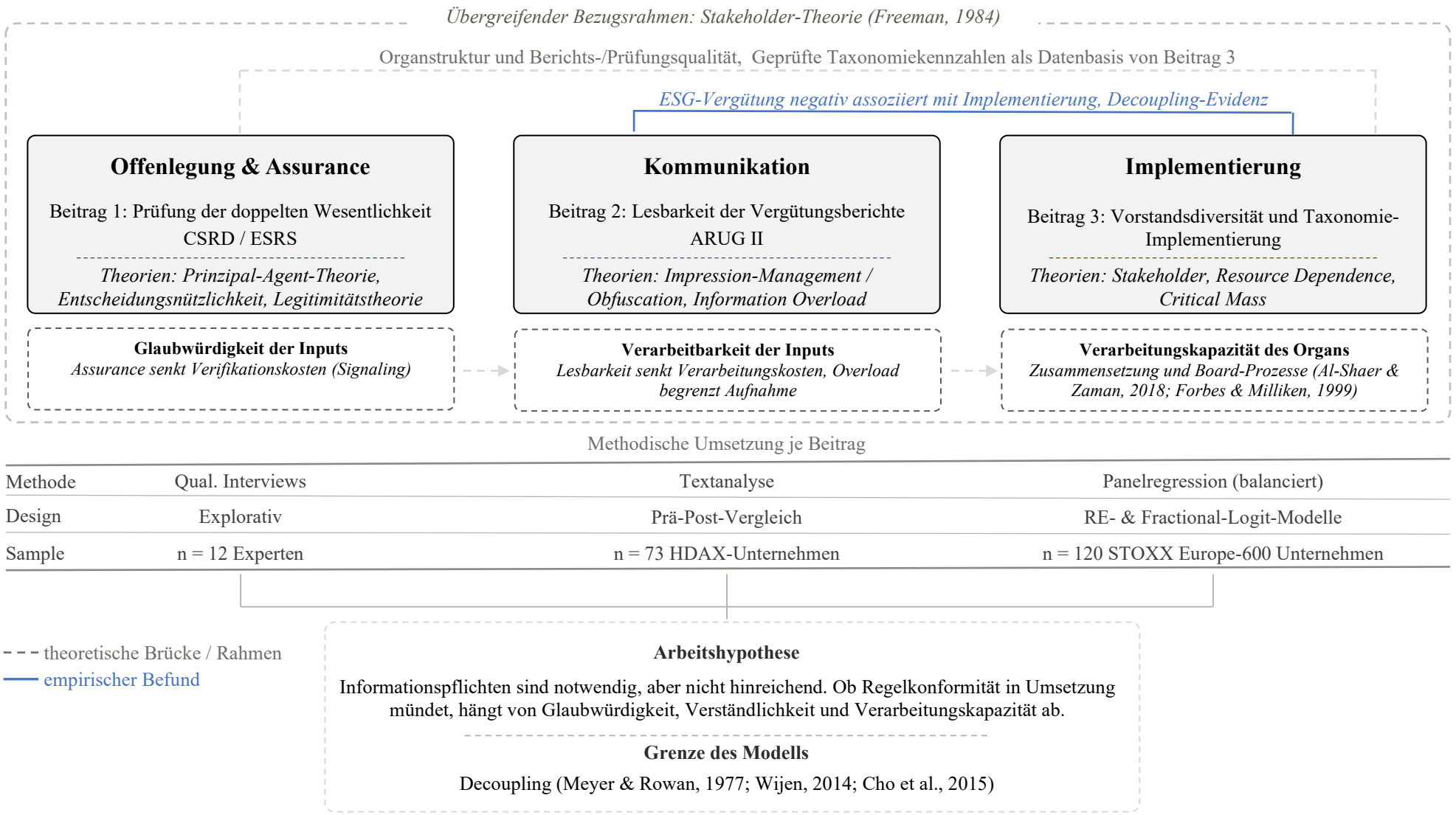


Abbildung 1.1: Konzeptioneller Rahmen

1.3.1 Die Prüfung der doppelten Wesentlichkeitsanalyse im Kontext der Nachhaltigkeitsberichterstattung

Der erste Beitrag der Dissertation widmet sich der Frage, wie sich die externe Prüfung der doppelten Wesentlichkeitsanalyse gemäß CSRD/ESRS in der Prüfungspraxis gestaltet und welche Chancen und Herausforderungen sich für Wirtschaftsprüfer sowie berichtspflichtige Unternehmen ergeben. Ausgangspunkt ist die durch die CSRD eingeführte Systematisierung der Wesentlichkeitsbestimmung. Anders als unter der NFRD, deren Wesentlichkeitsbegriff vage blieb und konkurrierende Auslegungen zuließ, die sich in den nationalen Umsetzungen unterschiedlich niederschlugen (Baumüller & Sopp, 2022; Sopp et al., 2022; Raith, 2023), verankert die CSRD mit den ESRS eine verbindliche Zweidimensionalität aus finanzieller und Impact-Wesentlichkeit, die zur prozessualen Schaltstelle der gesamten Nachhaltigkeitsberichterstattung wird (Miettinen, 2024). Dass die Bestimmung wesentlicher Themen dabei auch unter präziseren regulatorischen Vorgaben ermessensbehaftet bleibt, verdeutlichen Garst et al. (2022), die Wesentlichkeitsanalysen aufgrund der Komplexität, Unsicherheit und Bewertungsabhängigkeit von Nachhaltigkeitsthemen als kunstähnliche und nicht als rein wissenschaftliche Praxis charakterisieren. Der Forschungsstand zur ESG-Assurance zeigt zwei systematische Lücken, die auch für den vorliegenden Beitrag relevant sind. Erstens datiert der Großteil der empirischen Studien aus der Zeit vor der jüngsten regulatorischen Reform. Der aktuellste Überblicksbeitrag von Venter und Krasodomska (2024) berücksichtigt Literatur bis August 2023 und damit ausschließlich Studien, die vor Anwendung der CSRD entstanden sind. Zweitens konzentriert sich die Forschung stark auf große, börsennotierte Unternehmen in entwickelten Kapitalmärkten, während nicht börsennotierte Unternehmen, kleine und mittlere Unternehmen deutlich unterrepräsentiert bleiben (Alsahali & Malagueño, 2022; Hazaea et al., 2022). Eine erste Studie zur CSRD-Ära liegt inzwischen vor. Laub und Steller (2026) befragen siebzehn Wirtschaftsprüfer in Deutschland und Österreich und finden eine strukturelle Diskrepanz zwischen dem regulatorischen Anspruch der doppelten Wesentlichkeit und ihrer Anwendung in der Prüfungspraxis.

Die empirische Literatur zur ESG-Assurance im europäischen Kontext zeigt zudem, dass die NFRD zu einem messbaren Anstieg der Assurance-Praxis geführt hat. Ottenstein et al. (2022) dokumentieren anhand von 905 betroffenen Unternehmen aus 30 europäischen Ländern, dass diese rund 19 % häufiger eine externe Prüfung ihrer Nachhaltigkeitsberichterstattung erhalten als vergleichbare Kontrollunternehmen. Die Verbreitung fällt allerdings uneinheitlich aus. Vander Bauwhede & Van Cauwenberge (2022) identifizieren für 660 europäische Unternehmen im NFRD-Pflichtkontext Unternehmensgröße, ESG-Performance und Branchenzugehörigkeit als wesentliche Treiber der

Assurance-Entscheidung, während Simoni et al. (2020) auf länderspezifische institutionelle Faktoren verweisen.

Zugleich markiert die CSRD einen Bruch mit der bisherigen Prüfungspraxis, da sie erstmals eine verpflichtende externe Prüfung vorschreibt, zunächst mit begrenzter und perspektivisch mit hinreichender Sicherheit (Ruohonen & Kullas, 2024).[10] Umstritten ist dabei, wem diese Prüfung obliegen soll. Artikel 34 RL 2013/34/EU eröffnet den Mitgliedstaaten ein Wahlrecht, die materielle Prüfung durch den Abschlussprüfer, einen weiteren Wirtschaftsprüfer oder unabhängige Prüfungsdienstleister außerhalb des Berufsstandes zuzulassen. Der deutsche Regierungsentwurf sieht vorerst von diesem Wahlrecht ab und definiert die Prüfung gemäß § 324e HGB-E als Vorbehaltsaufgabe für Wirtschaftsprüfer. Ob diese Zuordnung sachgerecht ist, wird in der Literatur kontrovers diskutiert (Quick et al., 2022; Velte, 2022b).

Da die geprüften Berichte großer, nicht kapitalmarktorientierter Unternehmen erst ab 2026 vorliegen, wählt die Studie eine qualitative Herangehensweise. Befragt wurden zwölf Experten aus der Wirtschaftsprüfung, davon zwei aus Big-Four-Gesellschaften, acht aus Next-Ten-Gesellschaften und zwei aus kleineren Kanzleien, ausgewählt aus dreißig angefragten Personen. Die Erhebung erfolgte mittels leitfadengestützter, semistrukturierter Interviews, wobei der Leitfaden zuvor in einem Pre-Test validiert wurde (Mayring, 2023; Weichbold, 2022). Die Auswertung stützt sich auf die qualitative Inhaltsanalyse mit kombiniert deduktiver und induktiver Kategorienbildung nach dem sechsstufigen, MAXQDA-gestützten Verfahren von Kuckartz und Rädiker (2024).

Theoretisch stützt sich der Beitrag auf die Prinzipal-Agent-Theorie, wonach die externe Prüfung als Monitoring-Instrument zur Überwachung des Managements und zum Abbau von Informationsasymmetrien fungiert (Jensen & Meckling, 1976; Simnett et al., 2009; Quick, 2024), ergänzt um die Theorie der Entscheidungsnützlichkeit (Staubus, 1961; Staubus, 2013). Diese Theoriewahl korrespondiert mit dem Forschungsstand (Del Gesso und Lodhi, 2025). Empirisch untermauert wird die Argumentation dadurch, dass geprüfte Nachhaltigkeitsinformationen von Kapitalgebern signifikant höher bewertet werden als ungeprüfte (Clarkson et al., 2019; Reimsbach et al., 2018; García-Sánchez et al., 2019; Carey et al., 2021) und dass die Prüfung die Berichtsqualität erhöht, insbesondere wenn sie durch Wirtschaftsprüfer erfolgt (Zorio et al., 2013; Ballou et al., 2018).

[10] Diese regulatorisch vorgezeichnete Perspektive auf eine Anhebung des Sicherheitsniveaus wurde inzwischen aufgegeben. Die Richtlinie (EU) 2026/470 streicht die Ermächtigung der Kommission, Standards für Prüfungen mit hinreichender Sicherheit zu erlassen, sodass die Anforderung voraussichtlich bei der Prüfung mit begrenzter Sicherheit verbleibt.

Martínez-Ferrero et al. (2018) zeigen, dass Branchenspezialisierung und Prüfererfahrung die Qualität von Assurance-Berichten erhöhen und dieser Effekt zusätzlich verstärkt wird, wenn der Anbieter eine Wirtschaftsprüfungsgesellschaft ist. Ruiz-Barbadillo und Martínez-Ferrero (2020) weisen ergänzend nach, dass die gemeinsame Erbringung von Abschluss- und Nachhaltigkeitsprüfung durch denselben Anbieter Wissens-Spillover erzeugt und dadurch die Prüfungsqualität steigert. Velte (2025b) relativiert dieses Bild jedoch in einem strukturierten Überblick über 52 quantitative Studien, wonach die Befundlage zu heterogen ist, um klare Tendenzen hinsichtlich der Konsequenzen einer bestimmten Prüferwahl abzuleiten. Zengin-Karaibrahimoglu et al. (2026) führen diese Heterogenität auf zwei unterschiedliche Prüfungslogiken zurück. Wirtschaftsprüfungsgesellschaften folgen einem auditorientierten Modell, das auf interne Kontrollen, Prozessverifikation und finanzielle Wesentlichkeit abstellt, während spezialisierte Anbieter Stakeholderinklusion und qualitative Aspekte betonen. Für die Frage der Vorbehaltsaufgabe bedeutet das, dass mit der Prüferwahl zugleich eine Entscheidung über die Prüfungslogik getroffen wird, was die von den Befragten dieser Studie geäußerten Präferenzen in ein anderes Licht rückt. Insgesamt zeichnet sich damit das Bild eines sich rasch entwickelnden, zugleich aber weiterhin heterogenen Assurance-Markts, dessen Ausgestaltung eng mit regulatorischen Rahmenbedingungen, Corporate Governance Strukturen und kulturellen Kontexten verwoben ist (Kolk, 2008; Kolk & Perego, 2010; Perego & Kolk, 2012; Hazaea et al., 2022).

Fünf der zwölf Befragten sprechen sich explizit für die Vorbehaltsaufgabe der Wirtschaftsprüfer aus, begründet mit etablierten Qualitätssicherungs- und Haftungsregimen, während sich die übrigen nicht eindeutig positionieren. Zugleich wird die ergänzende Einbindung spezialisierter Expertise, etwa durch Umweltingenieure, analog zur Hinzuziehung von Versicherungsmathematikern bei Pensionsrückstellungen als sinnvoll erachtet. Besondere Synergieeffekte werden erwartet, wenn der Abschlussprüfer zugleich den Nachhaltigkeitsbericht prüft, was einer Entkopplung von Finanz- und Nachhaltigkeitsinformationen entgegenwirkt (so auch Velte & Stawinoga, 2022). Der Praxisstand erweist sich als heterogen. Zwar bestehen in den meisten Gesellschaften bereits Expertenteams, doch verfügen erst drei von zehn befragten Wirtschaftsprüfern über die IDW-Zusatzqualifikation Sustainability-Auditor. Die Einschätzungen zur Vorbehaltsaufgabe sind allerdings als Selbsteinschätzungen des eigenen Berufsstands zu lesen und insofern nicht frei von Eigeninteresse. Der Fachkräftemangel erschwert den angestrebten Personalaufbau, wird perspektivisch aber auch als Chance gesehen, das Nachwuchsproblem des Berufsstands zu adressieren (Velte, 2023).

Die Ergebnisse zeigen, dass die Prüfung der Wesentlichkeitsanalyse übereinstimmend als zentrales Instrument zur Sicherstellung von Vollständigkeit und Glaubwürdigkeit der Nachhaltigkeitsberichterstattung betrachtet wird und damit sowohl das Moral-Hazard-Problem im

Sinne der Prinzipal-Agent-Theorie (Jensen & Meckling, 1976) als auch das Risiko selektiv positiver Berichterstattung adressiert (Seele & Gatti, 2017). Die Vollständigkeit gilt über die Hälfte der Befragten als zentrale Assertion, da die größte Gefahr im absichtlichen oder datenbedingten Auslassen wesentlicher Themen bestehe.

Über den unmittelbaren Untersuchungsgegenstand hinaus verweist ein Interviewpartner darauf, dass leistungsabhängige Vergütungssysteme auf Basis von Nachhaltigkeitskennzahlen selbst Interessenkonflikte erzeugen können, wenn das Management primär auf die Maximierung der eigenen Vergütung ausgerichtet ist. Die Prüfung adressiert damit einen Zielkonflikt, der durch die Vergütungsarchitektur des Unternehmens erst hervorgebracht wird, und verbindet den ersten Beitrag sowohl mit der in Kapitel 3 untersuchten Vergütungsberichterstattung als auch mit Beitrag 3. Dort ist die Kopplung der Vorstandsvergütung an Nachhaltigkeitskennzahlen nicht positiv, sondern negativ mit der Taxonomie-Implementierung im Bereich der Betriebsausgaben assoziiert (Kapitel 4). Die aus der Prüfungspraxis geäußerte Vermutung findet damit eine erste, angesichts der begrenzten Stichprobe zurückhaltend zu interpretierende quantitative Entsprechung.

Als praktische Herausforderungen benennen die Befragten eine unzureichende Datenqualität, insbesondere entlang internationaler Lieferketten, sowie das Fehlen etablierter Standardisierung. Dieser Standardisierungsmangel wurde differenziert herausgearbeitet. Wesentlichkeitsschwellen seien sehr individuell, und es existiere kein einheitliches Vorgehen zur Überprüfung der zugrunde liegenden Kriterien und Quellen. Es fehle an normativen und greifbaren Zahlenwerken, auf deren Basis Entscheidungen getroffen werden könnten, weshalb stets ein Ermessensspielraum verbleibe, anders als bei quantitativen Sachverhalten mit klaren Fakten. Zwei Befragte kritisieren explizit, dass die daraus resultierenden Auslegungsfragen seitens der Regulierer nicht adressiert würden, und zwei weitere sehen die Gefahr, dass die Beurteilung derselben Wesentlichkeitsanalyse bei wechselndem Prüfer ceteris paribus zu unterschiedlichen Ergebnissen führt. Die Studie zeigt damit, dass das Fehlen einheitlicher Schwellenwerte nicht nur die Vergleichbarkeit der Berichte einschränkt, sondern unmittelbar auf die Prüfbarkeit durchschlägt.

Bereits Eccles et al. (2012) argumentieren, dass Wesentlichkeit und Berichtsstandards sektorweise entwickelt werden müssen, da ihr Fehlen zu inkonsistenten und teils irreführenden Offenlegungen führt. Baumüller und Sopp (2022) wiesen mit Blick auf die Vorläuferregulierung darauf hin, dass ökologische und soziale Wesentlichkeit weiterer Konzeptualisierung und Operationalisierung bedürfe, um Harmonisierung zu ermöglichen. Weshalb diese Forderung schwer einzulösen ist, verdeutlicht die eingangs referierte Charakterisierung von Wesentlichkeitsanalysen als ermessensabhängige Praxis (Garst et al., 2022). Spätere Arbeiten zeigen, dass die Lücke auch unter dem präziseren CSRD-Regime

fortbesteht. Miettinen (2024) dokumentiert für EU-Pharmaunternehmen, dass sich der Wesentlichkeitsbegriff zwar von einer vagen Konzeption unter der NFRD zur doppelten Wesentlichkeit unter der CSRD entwickelt hat, die Berichte jedoch weiterhin Inkonsistenzen aufweisen. Für die erste ESRS-Anwendung bestätigen Morganho et al. (2025) anhand von 45 iberischen Unternehmen erhebliche Lücken in den Wesentlichkeitsprozessen und deren Dokumentation, während Marín-Hernández et al. (2026) für den europäischen Wassersektor zeigen, dass eine breitere thematische Abdeckung nicht mit höherer Informationsqualität einhergeht. Für die Prüfungsperspektive dokumentieren Laub und Steller (2026) anhand von siebzehn Interviews in Deutschland und Österreich, dass das professionelle Urteil trotz formaler Verankerung der doppelten Wesentlichkeit in finanzieller Materialitätslogik verhaftet bleibt und sich die Prüfung auf vom Management definierte Prüfungsumfänge beschränkt.

Die Befragten nennen zwei Wege zur Schließung dieser Lücke. Der eine besteht in einer Konkretisierung durch Branchenstandards, Leitfäden und Praxishinweise, etwa vom Institut der Wirtschaftsprüfer (IDW) oder von Branchenverbänden. Der andere besteht in einer Angleichung der Prüfungspraxis, die sich im ersten Berichtsjahr aus dem fachlichen Austausch innerhalb des Berufsstands ergeben soll. Institutionentheoretisch beschreiben die Befragten damit normativen und mimetischen Isomorphismus als erwartete Ergänzung der koerziven Vorgabe (DiMaggio & Powell, 1983). Für den ersten Weg liegt inzwischen ein konkreter Vorschlag vor. Henkel et al. (2026) leiten auf Basis der DAX- und MDAX-Berichte für das Geschäftsjahr 2024 branchenspezifische quantitative Schwellenwerte aus Median, Quartilen und Interquartilsabständen der berichteten Kennzahlen ab, um die bislang überwiegend ordinal skalierten und damit schwer vergleichbaren Scoringverfahren der Impact-Wesentlichkeit zu objektivieren. Zugleich zeigen sie, dass die dafür erforderliche Datenbasis nur für einzelne Umwelt- und Sozialkennzahlen in belastbarer Qualität vorliegt, während andere ESRS-Themenfelder erhebliche Lücken aufweisen. Der Vorschlag markiert damit zugleich das Potenzial und die gegenwärtige Grenze einer datenbasierten Schwellenwertbestimmung.

Ob sich die von den Befragten erwartete Konvergenz tatsächlich einstellt, lässt sich mit den ersten Längsschnittdaten prüfen. Nahrstedt (2026) vergleicht die Wesentlichkeitsanalysen von 200 EU-Unternehmen zwischen den Geschäftsjahren 2024 und 2025 und dokumentiert eine Konsolidierung. Die Wesentlichkeitsquoten sanken für fünf der zehn themenbezogenen Standards, am deutlichsten für ESRS E2, während sie für ESRS S3 und G1 stiegen. Treiber der Konsolidierung war überwiegend die finanzielle Wesentlichkeitsdimension, und zwischen den Sektoren zeigen sich erhebliche Unterschiede. Dass die Anpassung vor allem über die finanzielle Dimension verläuft, korrespondiert mit der von Laub und Steller (2026) beschriebenen Dominanz finanzieller Materialitätslogik. Ob die

Konsolidierung Lernprozesse abbildet oder eine vorweggenommene Anpassung an das Omnibus-Paket, lässt sich anhand der Daten nicht trennen.

Hinsichtlich der Prüfungsplanung zeigt sich, dass die Prüfung der Wesentlichkeitsanalyse etwa ein Drittel der Gesamtprüfungszeit beansprucht, wobei der Aufwand im ersten Anwendungsjahr besonders hoch ausfällt und in den Folgejahren deutlich abnimmt. Mehrere Befragte bewerten eine frühzeitige Einbindung des Prüfers als vorteilhaft, da kritische Themen so rechtzeitig identifiziert und Gegenmaßnahmen ergriffen werden können. Als methodische Herangehensweise der Prüfungspraxis beschreibt ein Interviewpartner die Erstellung einer sogenannten Schatten-Wesentlichkeitsanalyse, die auf Unternehmenskenntnis und externen Stakeholder-Informationen beruht und als Erwartungsmaßstab dient. Der International Standard on Assurance Engagements (ISAE) 3000 (Revised) wird von mehreren Befragten als Ausweichstandard oder Notlösung bezeichnet, während zwei Interviewpartner vom Prüfungsstandard International Standard on Sustainability Assurance 5000 (ISSA 5000) mehr Klarheit hinsichtlich der Abgrenzung der Zusicherungsniveaus erwarten.[11] Bezüglich der Prüfungssicherheit besteht Einigkeit, dass sich beide Niveaus vor allem in Prüfungstiefe und Dokumentationsaufwand unterscheiden, wobei für hinreichende Sicherheit auf unzureichende Kontroll- und IT-Systeme auf Mandantenseite verwiesen wird.[12]

Für den Bestätigungsvermerk erwarten mehrere Befragte eine nicht unerhebliche Anzahl eingeschränkter Testate, unter anderem aufgrund fehlerhafter Wesentlichkeitsanalysen, was im Austausch mit BaFin und IDW bestätigt worden sei. Einigkeit besteht darin, dass Einschränkungen des Nachhaltigkeitsurteils keinen unmittelbaren Einfluss auf den finanziellen Bestätigungsvermerk haben, wobei durch den erweiterten Prognosezeitraum ein zeitversetzter Zusammenhang entstehen kann, da langfristige, über die doppelte Wesentlichkeit identifizierte Risiken sich erst später in der Finanzberichterstattung niederschlagen. Die erste Berichtssaison hat diese Erwartung bislang nicht bestätigt. Für das Geschäftsjahr 2024 dokumentieren Schmidt et al. (2025) anhand der Berichte deutscher Indexunternehmen, dass sämtliche geprüften Nachhaltigkeitsberichte ein uneingeschränktes Prüfungsurteil erhielten. Dieser Befund ist allerdings nur eingeschränkt aussagekräftig, da die Prüfung mangels rechtzeitiger CSRD-Umsetzung freiwillig erfolgte und damit einer Selbstselektion gut vorbereiteter Unternehmen unterlag, ganz überwiegend auf begrenzte Sicherheit lautete und sich auf börsennotierte Unternehmen mit vergleichsweise hoher Ressourcenausstattung beschränkte. Ob sich

[11] Der ISSA 5000 wurde im September 2024 durch das IAASB verabschiedet und ist für Prüfungszeiträume ab dem Geschäftsjahr 2026 anzuwenden (International Auditing and Assurance Standards Board (IAASB), 2024). Die zum Interviewzeitpunkt bestehende Unsicherheit über die Verabschiedung des Standards ist damit inzwischen ausgeräumt.

[12] Die von den Befragten erörterte Abgrenzung zwischen begrenzter und hinreichender Sicherheit hat durch die zwischenzeitliche Regulierungsentwicklung an praktischer Relevanz verloren.

die Erwartung der Befragten mit der verpflichtenden Prüfung eines breiteren Anwenderkreises einlöst, bleibt offen.

Als Nutzen der Prüfung heben die Befragten ein vertrauensbildendes Signal an Investoren, eine legitimierende Wirkung sowie potenziellen Einfluss auf Gesellschafterstruktur und Finanzierung hervor. Kritisch bewertet wird demgegenüber die fehlende Verhältnismäßigkeit der Berichtspflicht für kleinere Unternehmen mit vergleichsweise geringem Carbon Footprint. Diese zum Interviewzeitpunkt geäußerte Kritik wurde regulatorisch inzwischen aufgegriffen. Die im Rahmen des EU-Omnibus-Pakets verabschiedete Richtlinie (EU) 2026/470 hebt die Schwellenwerte der CSRD-Berichtspflicht auf mehr als 1.000 Mitarbeitende und mehr als 450 Millionen Euro Nettoumsatz an, wodurch je nach Schätzung rund 80 Prozent der ursprünglich berichtspflichtigen Unternehmen aus dem Anwendungsbereich herausfallen (Velte 2026). Die kritisierte Verhältnismäßigkeitsfrage ist damit inzwischen Gegenstand regulatorischer Nachsteuerung, nicht mehr nur offener Kritikpunkt aus der Prüfungspraxis.

1.3.2 Vergütungsberichterstattung der HDAX-Unternehmen nach ARUG II: Formale Aspekte und textanalytische Erkenntnisse

Wie in Kapitel 1.2 dargelegt, ist Corporate Governance in der hier entwickelten Informationsarchitektur nicht bloß Träger der Berichterstattung, sondern aktiver Hebel der Umsetzung. Vergütungssysteme stellen dabei einen zentralen Governance-Mechanismus dar, um Anreize zu setzen, Agency-Probleme zu begrenzen und Managementhandeln an langfristiger Unternehmensentwicklung auszurichten (Jensen & Murphy, 1990; Edmans & Gabaix, 2016). Im Zuge einer stärkeren Berücksichtigung von Stakeholderinteressen und Nachhaltigkeitszielen gewinnt diese Steuerungslogik zusätzlich an Bedeutung, da Unternehmen variable Vergütung zunehmend an ESG-Kriterien koppeln, um strategische Nachhaltigkeitsziele in Managemententscheidungen zu übersetzen (Flammer et al., 2019). Die Wirksamkeit dieser Steuerungslogik ist jedoch eng mit Transparenz verknüpft. Erst eine standardisierte und verständliche Offenlegung ermöglicht Monitoring und Vergleichbarkeit und stärkt damit die Rechenschafts- und Kontrollfunktion gegenüber Aktionären und anderen Interessensgruppen – entsprechend fordert die Richtlinie (EU) 2017/828 gem. Art. 9b ausdrücklich einen „klaren und verständlichen Vergütungsbericht ". Dass diese Verständlichkeit auch kapitalmarktseitig relevant ist, zeigt die Lesbarkeitsliteratur. Nach Mercer (2004) zählen die Merkmale der Offenlegung selbst zu einem zentralen Faktor, anhand derer Investoren die Glaubwürdigkeit einer Managementmitteilung beurteilen. Rennekamp (2012) weist experimentell nach, dass verständlichere Offenlegungen das Gefühl der Verarbeitungsflüssigkeit erhöhen und dadurch zu stärkeren Reaktionen insbesondere von Kleinanlegern führen, wobei sich die wahrgenommene Glaubwürdigkeit des Managements nicht unmittelbar, sondern nur vermittelt über diese Verarbeitungsflüssigkeit verändert. Asay et al. (2017) ergänzen, dass Investoren bei schlecht lesbaren Offenlegungen unsicherer in ihrer Bewertung werden und sich stärker auf externe Informationsquellen stützen. Tan et al. (2014) zeigen schließlich, dass die Wirkung sprachlicher Tonalität von der Lesbarkeit abhängt und sich je nach Erfahrungsgrad der Investoren gegenläufig entfalten kann.

Untersucht wird, wie sich Umfang und Lesbarkeit der Vergütungsberichte von HDAX-Unternehmen durch die Einführung des ARUG II verändert haben und inwiefern das in § 162 Abs. 1 Satz 1 AktG normierte Gebot eines klaren und verständlichen Berichts erfüllt wird. Der Beitrag adressiert damit eine explizite Forschungslücke, da weiterführende Studien zu den Eigenschaften des Vergütungsberichts als eigenständigem Berichtsinstrument im deutschsprachigen Raum bislang fehlen. Anhand eines Prä-Post-Vergleichs analysiert die Studie die Vergütungsberichte von 73 kontinuierlich im HDAX gelisteten Unternehmen für die Geschäftsjahre 2020 bis 2022 in deutscher und englischer Sprache, wobei 2020 die letzte Berichtsperiode vor und 2021 sowie 2022 die ersten

Perioden nach Inkrafttreten des Gesetzes abbilden. Die Sprachkomplexität wird über vier etablierte Indizes erfasst, für die deutschsprachigen Berichte über den Lesbarkeitsindex nach Björnsson (LIX; Björnsson, 1968, 1983) und den Flesch Reading Ease nach Amstad (FREAmstad; Amstad, 1978), für die englischsprachigen über den Flesch Reading Ease (FRE; Flesch, 1948) und den Gunning Fog Index (FOG; Gunning, 1952). Die Auswertung erfolgte über zwei eigens entwickelte, sprachspezifische Python-Programme, deren Ergebnisse ergänzend gegen webbasierte Werkzeuge validiert wurden.

Theoretisch stützt sich der Beitrag auf die Impression-Management- beziehungsweise Obfuscation-Theorie, wonach das Management die Berichterstattung so gestaltet, dass nachteilige Informationen sprachlich schwer identifizierbar bleiben (Courtis, 1998; Li, 2008; Brennan et al., 2009; Gosselin et al., 2021). Ergänzt wird dies um Überlegungen zu Zieldivergenzen zwischen Vorstand und Adressaten, die in zwei Konstellationen auftreten können. Zum einen kann der Vorstand an der Optimierung kurzfristiger variabler Vergütung interessiert sein, etwa durch den Ausschluss aufwandswirksamer Sachverhalte mittels Non-GAAP-Kennzahlen oder bilanzpolitischer Maßnahmen, während Aktionäre eine nachhaltige Wertsteigerung verfolgen, die typischerweise über langfristige variable Komponenten incentiviert wird (Li, 2008; Laksmana et al., 2012). Zum anderen können ESG-relevante Vergütungsziele lediglich symbolisch integriert und nicht ernsthaft verfolgt werden, womit Stakeholder-Interessen insgesamt vernachlässigt würden. Empirisch gestützt wird diese Vermutung durch Bannier und Reinschmidt (2022), wonach Nachhaltigkeitsziele eher in der kurzfristigen als in der langfristigen variablen Vergütung berücksichtigt werden, sowie durch internationale Befunde, wonach ESG-Kriterien in Vergütungsplänen überwiegend nicht bindend ausgestaltet sind (Bebchuk & Tallarita, 2022; Lu, 2023). Diese Konstellation verbindet den Beitrag unmittelbar mit der in Kapitel 1.2 diskutierten Decoupling-Problematik. Beitrag 3 liefert für diese Vermutung eigene Evidenz, denn dort ist die Kopplung der Vorstandsvergütung an Nachhaltigkeitskennzahlen nicht positiv, sondern negativ mit der Taxonomie-Implementierung assoziiert (Kapitel 1.3.3). Für die Beurteilung des gestiegenen Berichtsumfangs zieht der Beitrag zudem die Information-Overload-Theorie heran, wonach Informationsverarbeitungsanforderungen die verfügbare Zeit oder Kapazität der Adressaten übersteigen können und der Einsatz von Visualisierungen sowie Inhaltsverzeichnissen diese Belastung reduziert (Schick et al. 1990; Hartmann & Weißenberger, 2024).

Der Berichtsumfang hat deutlich zugenommen. Die durchschnittliche Wortzahl der deutschsprachigen Berichte stieg statistisch signifikant von 4.043 im Jahr 2020 auf 6.494 im Jahr 2021 und 6.666 im Jahr 2022, die der englischsprachigen von 5.205 auf 7.704 und 8.267. Der durchschnittliche Seitenumfang wuchs von 15,1 auf 27,75 und 30,11 Seiten. Diesem Anstieg steht eine nur unterproportionale Zunahme der visuellen Aufbereitung gegenüber. Zwar erhöhte sich die absolute Zahl der Abbildungen

und Tabellen je Bericht signifikant von 11,52 auf 19,64 und 21,05, ihre relative Dichte je Seite sank jedoch kontinuierlich von 0,77 über 0,74 auf 0,72. Der Anteil der Berichte mit Inhaltsverzeichnis stieg von 36,99 auf 49,32 Prozent.

Die Lesbarkeit hat sich marginal verschlechtert und verbleibt durchgehend auf schwierigem Niveau. Für die deutschsprachigen Berichte stieg der LIX-Wert statistisch signifikant von 65,39 im Jahr 2020 auf 67,44 im Jahr 2021 und lag 2022 bei 67,29, durchgehend oberhalb des Schwellenwerts von 60, ab dem Texte als sehr schwer verständlich gelten (Anderson, 1983; Courtis, 1995; Björnsson, 1983). Der FREAmstad-Wert sank entsprechend signifikant von 25,77 auf 23,22 und lag 2022 bei 23,66, womit die Berichte durchweg in die Referenzkategorie schwieriger Texte fallen (Amstad, 1978). Für die englischsprachigen Berichte zeigt sich dasselbe Muster. Der FRE-Wert fiel signifikant von 25,24 auf 22,60 und 22,44 und verbleibt damit konstant in der Kategorie sehr schwieriger Texte mit akademischem Anspruchsniveau, während der FOG-Wert signifikant von 20,26 auf 21,00 und 21,05 stieg, was dem Leselevel von Hochschulabsolventen entspricht (Flesch, 1948; Courtis, 1995). Da alle vier Indizes gleichgerichtet verlaufen, sind die Befunde als robust einzustufen.

Im internationalen Vergleich liegen die FOG-Werte auf ähnlichem Niveau wie in Studien zu US-amerikanischen Vergütungsberichten, für die Mittelwerte von 21,82 (Laksmana et al., 2012) und 23,6 (Hemmings et al., 2020) berichtet werden, während eine Untersuchung von FTSE-350-Unternehmen mit einem Mittelwert oberhalb von 17 auf bessere Lesbarkeit hindeutet (Hooghiemstra et al., 2017). Der über den Gesamtzeitraum ermittelte FOG-Mittelwert von 20,77 liegt deutlich über dem für geschäftliche Texte empfohlenen Höchstwert von zwölf (Laksmana et al., 2012). Gegenüber der Management Discussion and Analysis (MD&A), dem lageberichtähnlichen Berichtsteil US-amerikanischer Unternehmen (Lo et al., 2017) erweisen sich deutsche Vergütungsberichte damit als sprachlich anspruchsvoller. Der FRE-Mittelwert der englischsprachigen HDAX-Berichte liegt zwischen den Ergebnissen von Laksmana et al. (2012) mit 19,84 und dem von Velte (2021b) für HDAX-Unternehmen im Zeitraum 2014 bis 2018 ermittelten Wert von 26,64, wobei beide in die anspruchsvollste Lesbarkeitskategorie fallen. Für die deutschsprachigen Berichte fehlen vergleichbare textanalytische Untersuchungen mit LIX und FREAmstad, was die Relevanz der vorliegenden Studie im deutschsprachigen Raum unterstreicht.

Über die Lesbarkeit hinaus dokumentiert der Beitrag Heterogenität in Veröffentlichungspraxis und Prüfungstiefe. 43 der 73 Unternehmen integrierten den Vergütungsbericht in den Geschäftsbericht, in acht Fällen für 2021 und fünf Fällen für 2022 sogar in den Lagebericht, was der verbreiteten Rechtsauffassung einer notwendigen Trennung beider Instrumente widerspricht, da der Vergütungsbericht als Gemeinschaftsaufgabe von Vorstand und Aufsichtsrat entsteht, während die

Erstellung des Lageberichts gemäß § 264 Abs. 1 HGB allein dem Vorstand obliegt (Vetter & Tielmann, 2022; Kuthe & Reuters, 2022). Nur 54 Unternehmen führen die Berichte klar gekennzeichnet und chronologisch auf den Investor-Relations-Seiten, während sie bei 14 Unternehmen ausschließlich über archivierte Hauptversammlungsunterlagen zugänglich sind. Ob diese Praxis dem in Erwägungsgrund 32 der Richtlinie (EU) 2017/828 geforderten problemlosen Zugang genügt, erscheint fraglich, zumal Erwägungsgrund 33 der Richtlinie ausdrücklich auch potenziellen Anlegern und interessierten Akteuren eine Beurteilung der Vergütung ermöglichen soll. In beiden Berichtsjahren ließen jeweils 44 Unternehmen und damit rund 60 Prozent ihre Berichte über die nach § 162 Abs. 3 AktG vorgeschriebene formelle Mindestprüfung hinaus inhaltlich prüfen, was der Beitrag als bewusstes Transparenzsignal interpretiert, ohne dessen kapitalmarktbezogene Wirkung selbst zu untersuchen. Die vom IDW (2018) geäußerte Sorge vor einer wachsenden Erwartungslücke infolge einer rein formellen Prüfung erweist sich damit als unbegründet. Die Bewegung verläuft dabei allerdings in beide Richtungen, da sechs Unternehmen ihre Prüfungstiefe erhöhten und sechs andere sie reduzierten, jeweils ohne Angabe von Gründen. In zwei Fällen fehlt für das Berichtsjahr 2022 der gesetzlich vorgeschriebene Prüfvermerk vollständig, sodass externen Adressaten dort jede Form der Assurance fehlt.

Die dargestellten Befunde sind mit der eingangs skizzierten Governance-Logik vereinbar, wobei das deskriptive Prä-Post-Design ohne Kontrollgruppe keine Aussage darüber erlaubt, ob die beobachtete Entwicklung allein auf das ARUG II zurückgeht. Formale Regelkonformität in Gestalt von Berichtsumfang, Detailtiefe und Prüfungshäufigkeit hat spürbar zugenommen, doch die für eine wirksame Rechenschafts- und Kontrollfunktion entscheidende Verständlichkeit bleibt zurück. Die Anforderungen des § 162 AktG werden folglich nur bedingt erfüllt. Uneinheitliche Darstellungsformen und heterogene Veröffentlichungspraxis schränken zudem sowohl den intra- als auch den interunternehmerischen Vergleich ein. Das stützt die in Kapitel 1.2 entwickelte These, dass regulatorische Informationspflichten notwendig, aber nicht per se hinreichend für die intendierte Wirkung sind.

1.3.3 Executive Board Gender Diversity and ESG Performance: Evidence from EU Taxonomy Disclosures

Der dritte Beitrag untersucht empirisch, ob die Geschlechterzusammensetzung von Leitungsgremien börsennotierter Unternehmen mit der tatsächlichen Umsetzung der EU-Taxonomie-Anforderungen zusammenhängt – nicht nur mit deren Offenlegung. Die bisherige CSR-/ESG-Forschung hat den Zusammenhang zwischen Board-Struktur, Nachhaltigkeitsberichterstattung und CSR-Performance vielfach untersucht. Weitgehend offen blieb jedoch, welche Governance-Merkmale tatsächlich mit der materiellen Implementierung regulatorischer Nachhaltigkeitsvorgaben korrespondieren, also mit dem, was über die reine Offenlegung hinausgeht (Campopiano et al., 2023; Chrzan & Pott, 2024; Nicolò et al., 2022). An dieser Lücke setzt der Beitrag an und fragt, ob und unter welchen Bedingungen ein höherer Frauenanteil im Vorstand (Executive Board Gender Diversity, BGD) mit einer stärkeren Umsetzung der EU-Taxonomie einhergeht.

Den theoretischen Rahmen bilden drei sich ergänzende Ansätze. Die Stakeholder-Theorie liefert die Grundannahme, dass diverser besetzte Gremien gesellschaftliche und regulatorische Nachhaltigkeitserwartungen aufmerksamer aufgreifen (Freeman, 1984; Michelon & Parbonetti, 2012). Die Resource-Dependence-Theorie erklärt dies über die Ressourcenfunktion von Boards, wonach Vorstandsmitglieder Expertise, Netzwerke und Legitimität als kritische Ressourcen einbringen (Hillman & Dalziel, 2003; Pfeffer & Salancik, 1978). Die Critical Mass-Theorie präzisiert schließlich, dass Diversität typischerweise erst ab etwa drei weiblichen Mitgliedern beziehungsweise einem Anteil von rund 30 Prozent Wirkung entfaltet und nicht bereits bei einzelner, tokenhafter Repräsentation (Kanter, 1977; De Masi et al., 2021; Velte, 2025a). Diese Schwellenannahme ist in der jüngeren Literatur allerdings nicht unbestritten (Velte, 2025a). Der Beitrag prüft die Schwelle daher empirisch, statt sie vorauszusetzen.

Konzeptionell schließt der Beitrag an die fortschreitende Standardisierung nachhaltigkeitsbezogener Kennzahlen an, wie sie die EU-Taxonomie-Verordnung gemeinsam mit CSRD und SFDR vorantreibt (Europäische Kommission, 2018; Bebbington et al., 2023). Mit den Kennzahlen nach Art. 8 der Verordnung (EU) 2020/852 in Verbindung mit der Delegierten Verordnung (EU) 2021/2178 lassen sich Nachhaltigkeitsziele über standardisierte und prüfungspflichtige Größen abbilden, konkret den taxonomiekonformen Anteil von Umsatz, Investitionsausgaben (CapEx) und Betriebsausgaben (OpEx). Ob diese Kennzahlen die intendierte Vergleichbarkeit tatsächlich herstellen, ist Gegenstand laufender Diskussion (Nipper et al., 2025; Schütze & Sandbaek, 2025). Der Beitrag argumentiert jedoch, dass standardisierte Kennzahlen allein noch keine substanzielle Verhaltensänderung bewirken. Entscheidend sind vielmehr die Governance-Strukturen und Entscheidungsprozesse, die diese

Kennzahlen erst in Handeln übersetzen (Adams, 2002, 2004; Hahn & Kühnen, 2013; Aguilera et al., 2021; Bebbington et al., 2023). Executive BGD wird entsprechend als ein solcher Übersetzungsmechanismus verstanden, der regulatorische Anforderungen in konkrete Investitions-, Allokations- und Umsetzungsentscheidungen überführt.

Empirische Grundlage ist ein Panel von 120 STOXX-Europe 600-Unternehmen aus 13 Ländern und 23 Branchen über die Jahre 2022 bis 2024 (360 Firmenjahresbeobachtungen). Um Verzerrungen durch Schätzwerte auszuschließen, fließen ausschließlich selbst berichtete ("as reported") Taxonomie-Kennzahlen ein. Vier eigens operationalisierte Implementierungsraten dienen als abhängige Variablen, aggregiert sowie getrennt nach Umsatz-, CapEx- und OpEx-Konformität. Die Geschlechterdiversität im Vorstand wird sowohl als kontinuierliches Maß als auch als Critical Mass-Dummy (Frauenanteil ≥ 30 %) erfasst. Kontrolliert wird für finanzielle Merkmale (Eigenkapitalrendite, Verschuldungsgrad, Unternehmensgröße), governancebezogene Faktoren (CSR-Ausschuss, Board-Unabhängigkeit, nachhaltigkeitsbezogene Vergütung) sowie Zeit- und Brancheneffekte. Geschätzt wird primär mit Random-Effects-Panelregressionen und zweifach clusterrobusten Standardfehlern, ergänzt um Fractional-Logit-Modelle zur Berücksichtigung der begrenzten Wertespanne der Implementierungsraten (Papke & Wooldridge, 1996; Petersen, 2009; Wooldridge, 2010; Cameron et al., 2011; Gallani et al., 2015)

Der positive Zusammenhang zwischen Executive BGD und Taxonomie-Implementierung zeigt sich konsistent über die Modelle hinweg und verstärkt sich deutlich, sobald die 30-Prozent-Schwelle erreicht ist. Am robustesten fällt der Effekt für CapEx- und OpEx-basierte Implementierungsraten sowie für die aggregierte Kennzahl aus. Für die umsatzbasierte Konformitätsrate ist die Evidenz schwächer. Plausibel erscheint das, weil CapEx- und OpEx-Entscheidungen unmittelbarer im Einflussbereich der Unternehmensleitung liegen, während der Umsatzanteil taxonomiekonformer Aktivitäten stärker von Geschäftsmodell und Marktumfeld vorgegeben ist (Chrzan & Pott, 2024; Schütze & Sandbaek, 2025). Robustheitsprüfungen mit Fractional-Logit-Modellen sowie Interaktionstests mit der CO_2-Intensität der Unternehmen bestätigen dieses Bild. Der Diversitätseffekt bleibt unabhängig von der CO_2-Intensität bestehen, die ihrerseits primär das allgemeine Umsetzungsniveau beeinflusst. Verzögerte Spezifikationen deuten zudem auf ein Muster hin, das mit einer zeitlichen Abfolge vereinbar wäre, bei der sich Diversitätseffekte zunächst operativ, anschließend in Investitionsentscheidungen und erst mit weiterer Verzögerung im Umsatz niederschlagen. Angesichts des kurzen Beobachtungszeitraums von drei Jahren erlaubt dieser Befund allerdings keine belastbare Aussage über kausale Sequenzen.

Zwei Kontrollvariablen weisen zudem Vorzeichen auf, die der üblichen Erwartung zuwiderlaufen. Das Vorhandensein eines CSR-Ausschusses ist durchgängig negativ mit der Implementierung assoziiert und erreicht für OpEx Signifikanz, und die Kopplung der Vorstandsvergütung an Nachhaltigkeitskennzahlen ist im Fractional-Logit-Modell für OpEx auf dem Fünfprozentniveau negativ signifikant. Beide Merkmale gelten gemeinhin als Signal für Nachhaltigkeitsorientierung und sollten daher positiv mit der Umsetzung korrelieren. Eine Deutung liegt darin, dass diese Instrumente eher als Legitimationsvorrichtungen denn als Steuerungsmechanismen wirken. Das deckt sich mit Befunden, wonach ESG-Kriterien in Vergütungssystemen häufig unverbindlich ausgestaltet sind, nur geringes Gewicht in der Gesamtvergütung haben und sich für eine symbolische Übernahme eignen (Maas et al., 2016; Haque & Ntim, 2020; Bebchuk & Tallarita, 2022; Lu, 2023). Dass rund 90 Prozent der untersuchten Unternehmen einen CSR-Ausschuss unterhalten, begrenzt zudem die Unterscheidungskraft dieses Merkmals und lässt sich als Hinweis darauf lesen, dass seine Einrichtung eher dem Angleichungsdruck im organisationalen Feld folgt als einer unternehmensindividuellen Steuerungsentscheidung (DiMaggio & Powell, 1983). Angesichts der begrenzten Stichprobe und der Konzentration der Beobachtungen an den Verteilungsrändern sind diese Koeffizienten zurückhaltend zu interpretieren. Das Muster verdient gleichwohl Beachtung, weil es nahelegt, dass formal institutionalisierte Nachhaltigkeitsgovernance für sich genommen nicht mit höherer Umsetzung einhergeht, während die personelle Zusammensetzung des Vorstands es tut. Für die Informationsarchitektur der Dissertation ist dieser Befund von besonderem Wert, weil er die Vergütungsstruktur, deren Ausrichtung auf nachhaltige Unternehmensentwicklung das ARUG II vorgibt, mit der Umsetzung auf Ebene 3 verknüpft. Er stellt damit eine Verbindung zwischen einem der untersuchten Regelwerke und der Implementierungsebene her, nicht zwischen den Ebenen 2 und 3 im Sinne des Modells, da die Kommunikationsqualität des Vergütungsberichts selbst in Beitrag 3 nicht erfasst wird.

Theoretisch erweitert der Beitrag eine Literatur, die sich bislang schwerpunktmäßig mit Berichtsqualität, Transparenz und Reputation befasst hat (Tamimi & Sebastianelli, 2017; Arayssi et al., 2020; Nicolò et al., 2022), um einen untersuchten Zusammenhang, der Geschlechterdiversität als Kanal versteht, über den Governance-Strukturen mit tatsächlicher Umsetzung und nicht nur mit Berichterstattung in Verbindung stehen. Stakeholder-theoretisch gelesen spricht dies dafür, dass diverse Gremien heterogene Stakeholder-Erwartungen wirksamer in strategisches Handeln übersetzen. Ressourcenabhängigkeitstheoretisch gelesen erweitern weibliche Vorstandsmitglieder die organisationale Basis an Legitimität, Netzwerken und nachhaltigkeitsrelevanter Expertise und stärken damit die Anpassungsfähigkeit an regulatorische Anforderungen. Der Beitrag benennt zugleich zwei

Vorbehalte gegenüber dieser Deutung. Zum einen könnte der stark technische und compliancegetriebene Charakter der Taxonomie-Berichterstattung den Einflussspielraum des Leitungsgremiums begrenzen und die tatsächliche Steuerung auf spezialisierte Nachhaltigkeits-, Rechts- und Finanzfunktionen verlagern. Zum anderen unterliegen Unternehmen mit hoher Taxonomie-Konformität möglicherweise bereits einem so ausgeprägten regulatorischen und marktseitigen Druck, dass die Umsetzung eher externen Zwängen als internen Governance-Merkmalen folgt. Der Zusammenhang zwischen Vorstandszusammensetzung und Taxonomie-Implementierung ist damit nicht eindeutig determiniert.

Innerhalb der übergreifenden Drei-Ebenen-Architektur der Dissertation (Kap. 1.2) übernimmt der dritte Beitrag die Implementierungsebene und ergänzt damit die ersten beiden Beiträge komplementär. Während Beitrag 1 die Prüfung der doppelten Wesentlichkeitsanalyse nach CSRD/ESRS als Glaubwürdigkeitsmechanismus auf der Ebene von Offenlegung und Assurance untersucht und Beitrag 2 die Kommunikationsqualität der Vergütungsberichterstattung nach ARUG II analysiert, verschiebt der dritte Beitrag den Fokus auf akteurs- und strukturbezogene Governance-Mechanismen der obersten Unternehmensleitungsebene. Empirisch zeigt der Beitrag, dass die Umsetzung standardisierter Nachhaltigkeitsregulierung mit der personellen Zusammensetzung der Leitungsorgane zusammenhängt und nicht allein von der Existenz von Offenlegungs- und Prüfungspflichten bestimmt wird. Damit fügt der dritte Beitrag der Informationsarchitektur der Dissertation eine strukturelle, entscheidungsbasierte Dimension hinzu. Offenlegung und Assurance, Kommunikationsqualität und personelle Governance-Struktur lassen sich damit als drei Voraussetzungen verstehen, von deren Zusammenspiel abhängen dürfte, ob regulatorische Informationspflichten in reale unternehmerische Transformation münden. Ob dieses Zusammenspiel tatsächlich in der vermuteten Weise wirkt, bleibt empirisch offen.

1.4 Gemeinsamer Beitrag

In ihrer Gesamtheit stützen die drei Beiträge die Einschätzung, dass die Wirksamkeit der europäischen Nachhaltigkeitsregulierung wesentlich davon abhängt, wie Berichterstattung, Prüfung und Governance ineinandergreifen. Diese Einschätzung ist konzeptioneller Natur, da keiner der Beiträge das Zusammenwirken der drei Ebenen selbst testet. Während die doppelte Wesentlichkeit gem. ESRS 1.4 einen breiten Stakeholderansatz institutionalisiert, sorgen Standardisierung und Assurance für Glaubwürdigkeit und Vergleichbarkeit der Informationen. Governance wirkt schließlich als Übersetzungsinstanz, über die regulatorische Anforderungen in konkrete Investitions- und Transformationsentscheidungen überführt werden. Damit leistet die Arbeit einen integrierten Beitrag zur Forschung an der Schnittstelle von Nachhaltigkeitsberichterstattung, Corporate Governance und Sustainable Finance und liefert zugleich praxis- und politikrelevante Implikationen für Unternehmen, Prüfer und Regulierungsinstitutionen.

Die bisherige Forschung zur Wirkung verpflichtender Nachhaltigkeitsberichterstattung ist überwiegend fragmentiert. Christensen et al. (2021) zeigen, dass sich ökonomische Konsequenzen von Offenlegungspflichten nur unter Berücksichtigung von Ausgestaltung und Durchsetzung verstehen lassen, ohne dass ihre Synthese die drei hier untersuchten Ebenen systematisch verknüpft. Auch Hummel und Jobst (2024) beschreiben CSRD, Taxonomie und SFDR zwar als zunehmend interdependente Regulierungsarchitektur, bleiben dabei aber auf der Ebene der Rechtsakte selbst, ohne die dahinterliegenden Governance-Mechanismen zu adressieren, die diese Architektur erst wirksam werden lassen. Genau hier setzt die vorliegende Arbeit an. Sie untersucht mit der Glaubwürdigkeit auf der Ebene von Offenlegung und Assurance, der Verständlichkeit auf der Ebene der Kommunikationsqualität sowie den personellen Voraussetzungen der Umsetzung auf der Implementierungsebene drei Qualitätsdimensionen, die in der Literatur bislang überwiegend getrennt betrachtet werden. Die Befunde der drei Beiträge legen nahe, dass diese Dimensionen nicht isoliert wirken, sondern als ineinandergreifende Voraussetzungen dafür zu verstehen sind, dass Regulierung reale Wirkung entfaltet. Diese Zusammenschau leistet die Arbeit konzeptionell, während die empirische Prüfung des Zusammenwirkens Gegenstand künftiger Forschung bleibt.

Zugleich liefern die Beiträge aus drei Perspektiven Indizien, die mit der in Kapitel 1.2 entwickelten These vereinbar sind. Krueger et al. (2024) zeigen global, dass Offenlegungspflichten vor allem unter strenger Durchsetzung wirken. Die vorliegende Arbeit legt nahe, dass Durchsetzung selbst mehrdimensional zu denken ist und externe Prüfung,

verständliche Kommunikation sowie eine Governance-Struktur umfasst, die regulatorische Signale in Entscheidungen zu übersetzen vermag. Bartosch et al. (2026) ergänzen diese Durchsetzungsdimensionen um einen weiteren Kanal. Sie zeigen für die NFRD, dass verpflichtende Offenlegung die Zahl stakeholderseitiger Anschuldigungen gegen berichtende Unternehmen erhöht hat und dass umfangreichere Berichterstattung mit stärkerer Kritik einhergeht. Offenlegungspflichten wirken demnach nicht nur über Kapitalmärkte, sondern auch dadurch, dass sie Stakeholdern die Grundlage für Rechenschaftsforderungen liefern. Damit schließt die Arbeit an eine disziplinübergreifend geführte Debatte an, in der Ameli et al. (2020) für institutionelle Investoren bereits gezeigt haben, dass Transparenz allein keine hinreichende Antwort auf Klimarisiken darstellt. Die Decoupling-Literatur (Meyer & Rowan, 1977; Wijen, 2014) mahnt dabei zur Vorsicht, denn auch die hier dokumentierten positiven Zusammenhänge, etwa zwischen Board-Diversität und Taxonomie-Implementierung, belegen Assoziation und nicht automatisch vollständige Zielerreichung. Dass verpflichtende Berichterstattung Decoupling zumindest verringern kann, zeigen Aboud et al. (2024), ohne dass sie es beseitigt.

Für die Praxis ergeben sich drei Ansatzpunkte. Beitrag 1 legt nahe, dass eine frühzeitige Einbindung der Prüfer sowie eine Konkretisierung durch Branchenstandards und berufsständische Praxishinweise die dokumentierten Standardisierungsdefizite adressieren könnten. Beitrag 2 zeigt, dass regulatorisch geforderte Verständlichkeit nicht automatisch mit zunehmendem Berichtsumfang erreicht wird, was für die Ausgestaltung künftiger Berichtsformate unmittelbar relevant ist. Beitrag 3 liefert Aufsichtsräten und Nominierungsausschüssen ein Argument dafür, Geschlechterdiversität nicht allein als Compliance-Ziel, sondern als Bestandteil tatsächlicher Umsetzungsfähigkeit zu behandeln.

Die zentrale Grenze der Arbeit wurde bereits in Kapitel 1.2 benannt und betrifft die fehlende empirische Verkettung zwischen den drei Ebenen. Keiner der Beiträge testet direkt, ob höhere Prüfungsqualität zu verständlicheren Berichten oder verständlichere Berichterstattung zu höherer Umsetzungsfähigkeit führt. Die Beiträge belegen die Bausteine der Informationsarchitektur, nicht deren kausale Verkettung im Zeitverlauf.

Dass die Verkettung zwischen den Ebenen tatsächlich fragil ist, bestätigt sich bereits im ersten CSRD-Berichtsjahr. Vishnu Nampoothiri und Kumar (2026) dokumentieren anhand von 100 Assurance-Berichten aus 29 Ländern eine strukturierte Fragmentierung, denn 87 Prozent der Unternehmen erhalten zwar formal einheitlich benannte begrenzte Prüfungssicherheit, diese beruht jedoch auf grundlegend unterschiedlichen nationalen Rechtsrahmen, was die Autoren als legitimiertes Decoupling durch sanktionierte nationale Umsetzungspfade einordnen. Ein

analoges Muster zeigt sich auf der Berichtsebene. Pacelli und Ekinci (2026) kommen in einer vergleichenden Inhaltsanalyse von zehn Unternehmen aus verschiedenen Branchen zu dem Ergebnis, dass sich die Struktur der ESRS-Berichterstattung zwar zunehmend vereinheitlicht, die Anwendungsstrenge zwischen den Unternehmen jedoch erheblich variiert, sodass zwingende Regulierung die Form der Offenlegung standardisieren kann, ohne substanzielle organisationale Veränderung zu gewährleisten. Hadro et al. (2026) belegen dies für 53 europäische Automobilunternehmen anhand des Standards ESRS E5, wo symbolische Compliance-Rhetorik die substanzielle Auseinandersetzung mit zirkulären Strategien deutlich überwiegt.

Diese Einschränkungen markieren zugleich den Ansatzpunkt künftiger Forschung. Eine längsschnittliche Untersuchung, die Prüfungsqualität, Berichtsverständlichkeit und Gremienzusammensetzung über mehrere Jahre am selben Unternehmenssample verfolgt, könnte die hier konzeptionell hergestellte Verkettung empirisch prüfen. Ebenso offen bleibt, wie sich die Umsetzungsqualität unter dem im Anwendungsbereich reduzierten Omnibus-Regime entwickelt (Sabbatucci et al., 2026). Die zentrale Frage dieser Arbeit stellt sich damit unter veränderten regulatorischen Vorzeichen erneut, denn auch eine vereinfachte Informationsarchitektur bleibt in ihrer Wirksamkeit von der Qualität ihrer Umsetzung abhängig.

Literaturverzeichnis

Aboud, Ahmed; Saleh, Ahmed; Eliwa, Yasser (2024): Does mandating ESG reporting reduce ESG decoupling? Evidence from the European Union's Directive 2014/95. In: Business Strategy and the Environment 33 (2), S. 1305–1320. DOI: 10.1002/bse.3543.

Adams, Carol A. (2002): Internal organisational factors influencing corporate social and ethical reporting. In: Accounting, Auditing & Accountability Journal 15 (2), S. 223–250. DOI: 10.1108/09513570210418905.

Adams, Carol A. (2004): The ethical, social and environmental reporting-performance portrayal gap. In: Accounting, Auditing & Accountability Journal 17 (5), S. 731–757. DOI: 10.1108/09513570410567791.

Aguilera, Ruth V.; Aragón-Correa, J. Alberto; Marano, Valentina; Tashman, Peter A. (2021): The Corporate Governance of Environmental Sustainability: A Review and Proposal for More Integrated Research. In: Journal of Management 47 (6), S. 1468–1497. DOI: 10.1177/0149206321991212.

Aguilera, Ruth V.; Williams, Cynthia A.; Conley, John M.; Rupp, Deborah E. (2006): Corporate Governance and Social Responsibility: a comparative analysis of the UK and the US*. In: Corporate Governance 14 (3), S. 147–158. DOI: 10.1111/j.1467-8683.2006.00495.x.

Al-Shaer, Habiba; Zaman, Mahbub (2016): Board gender diversity and sustainability reporting quality. In: Journal of Contemporary Accounting & Economics 12 (3), S. 210–222. DOI: 10.1016/j.jcae.2016.09.001.

Al-Shaer, Habiba; Zaman, Mahbub (2018): Credibility of sustainability reports: The contribution of audit committees. In: Business Strategy and the Environment 27 (7), S. 973–986. DOI: 10.1002/bse.2046.

Alkaraan, Fadi; Elmarzouky, Mahmoud; Hussainey, Khaled; Venkatesh, V. G. (2023): Sustainable strategic investment decision-making practices in UK companies: The influence of governance mechanisms on synergy between industry 4.0 and circular economy. In: Technological Forecasting and Social Change 187, S. 122187. DOI: 10.1016/j.techfore.2022.122187.

Alliance for Corporate Transparency (2019): 2019 Research Report. An analysis of the sustainability reports of 1000 companies pursuant to the EU Non-Financial Reporting Directive. Online verfügbar unter https://www.allianceforcorporatetransparency.org/.

Alliance for Corporate Transparency (2020): 2020 Research Report. An analysis of the climate-related disclosures of 300 companies from Central, Eastern and Southern Europe pursuant to the EU Non-Financial Reporting Directive. Online verfügbar unter https://www.allianceforcorporatetransparency.org/.

Alsahali, Kholod Fahad; Malagueño, Ricardo (2022): An empirical study of sustainability reporting assurance: current trends and new insights. In: Journal of Accounting & Organizational Change 18 (5), S. 617–642. DOI: 10.1108/jaoc-05-2020-0060.

Ameli, Nadia; Drummond, Paul; Bisaro, Alexander; Grubb, Michael; Chenet, Hugues (2020): Climate finance and disclosure for institutional investors: why transparency is not enough. In: Climatic Change 160 (4), S. 565–589. DOI: 10.1007/s10584-019-02542-2.

Amstad, Toni (1978): Wie verständlich sind unsere Zeitungen? Dissertation. Zürich: Universität Zürich.

Anderson, Jonathan (1983): Lix and Rix. Variations on a Little-Known Readability Index. In: Journal of Reading 26 (6), S. 490–496.

Arayssi, Mahmoud; Jizi, Mohammad; Tabaja, Hala Hussein (2020): The impact of board composition on the level of ESG disclosures in GCC countries. In: Sustainability Accounting, Management and Policy Journal 11 (1), S. 137–161. DOI: 10.1108/SAMPJ-05-2018-0136.

Armstrong McKay, David I.; Staal, Arie; Abrams, Jesse F.; Winkelmann, Ricarda; Sakschewski, Boris; Loriani, Sina et al. (2022): Exceeding 1.5°C global warming could trigger multiple climate tipping points. In: Science 377 (6611), eabn7950. DOI: 10.1126/science.abn7950.

Asay, H. Scott; Elliott, W. Brooke; Rennekamp, Kristina (2017): Disclosure Readability and the Sensitivity of Investors' Valuation Judgments to Outside Information. In: The Accounting Review 92 (4), S. 1–25. DOI: 10.2308/accr-51570.

Bais, Beatrice; Nassimbeni, Guido; Orzes, Guido (2024): Global Reporting Initiative: Literature review and research directions. In: Journal of Cleaner Production 471, S. 143428. DOI: 10.1016/j.jclepro.2024.143428.

Ballou, Brian; Chen, Po-Chang; Grenier, Jonathan H.; Heitger, Dan L. (2018): Corporate social responsibility assurance and reporting quality: Evidence from restatements. In: Journal of Accounting and Public Policy 37 (2), S. 167–188. DOI: 10.1016/j.jaccpubpol.2018.02.001.

Bannier, Christina E.; Reinschmidt, Julia (2022): Vorstandsvergütung in DAX- und MDAX-Unternehmen. Wie nachhaltig sind die Vergütungssysteme der börsennotierten deutschen Unternehmen? In: WPg – Die Wirtschaftsprüfung 75, S. 639–645.

Bartosch, Julia; Avetisyan, Emma; Thompson, Paul A. (2026): Government Orchestration at Play: The Influence of Disclosure Regulation on Stakeholder Accusations and the Role of CSR. In: Journal of Management Studies. DOI: 10.1111/joms.70130.

Battiston, Stefano; Monasterolo, Irene; Riahi, Keywan; van Ruijven, Bas J. (2021): Accounting for finance is key for climate mitigation pathways. In: Science 372 (6545), S. 918–920. DOI: 10.1126/science.abf3877.

Baumüller, Josef; Sopp, Karina (2022): Double materiality and the shift from non-financial to European sustainability reporting: review, outlook and implications. In: Journal of Applied Accounting Research 23 (1), S. 8–28. DOI: 10.1108/jaar-04-2021-0114.

Baur, Nina; Blasius, Jörg (Hg.) (2022): Handbuch Methoden der empirischen Sozialforschung. Wiesbaden: Springer Fachmedien Wiesbaden.

Bebbington, Jan; Laine, Matias; Larrinaga, Carlos; Michelon, Giovanna (2023): Environmental Accounting in the European Accounting Review : A Reflection. In: European Accounting Review 32 (5), S. 1107–1128. DOI: 10.1080/09638180.2023.2254351.

Bebchuk, Lucian A.; Tallarita, Roberto (2022): The Perils and Questionable Promise of ESG-Based Compensation. In: Journal of Corporation Law 48 (1), S. 37–75. DOI: 10.2139/ssrn.4048003.

Berg, Florian; Kölbel, Julian F.; Rigobon, Roberto (2022): Aggregate Confusion: The Divergence of ESG Ratings. In: Review of Finance 26 (6), S. 1315–1344. DOI: 10.1093/rof/rfac033.

Björnsson, Carl Hugo (1968): Läsbarhet. Stockholm: Liber.

Björnsson, Carl Hugo (1983): Readability of Newspapers in 11 Languages. In: Reading Research Quarterly 18 (4), S. 480–497.

Boiral, Olivier; Henri, Jean-François (2017): Is Sustainability Performance Comparable? A Study of GRI Reports of Mining Organizations. In: Business & Society 56 (2), S. 283–317. DOI: 10.1177/0007650315576134.

Boiral, Olivier; Heras-Saizarbitoria, Iñaki (2020): Sustainability reporting assurance: Creating stakeholder accountability through hyperreality? In: Journal of Cleaner Production 243, S. 118596. DOI: 10.1016/j.jclepro.2019.118596.

Brennan, Niamh M.; Guillamon-Saorin, Encarna; Pierce, Aileen (2009): Impression management. Developing and illustrating a scheme of analysis for narrative disclosures. In: Accounting, Auditing & Accountability Journal 22 (5), S. 789–832. DOI: 10.1108/09513570910966379.

Bua, Giovanna; Kapp, Daniel; Ramella, Federico; Rognone, Lavinia (2024): Transition versus physical climate risk pricing in European financial markets: a text-based approach. In: The European Journal of Finance 30 (17), S. 2076–2110. DOI: 10.1080/1351847X.2024.2355103.

Buchetti, Bruno; Arduino, Francesca Romana; Perdichizzi, Salvatore (2025): A literature review on corporate governance and ESG research: Emerging trends and future directions. In: International Review of Financial Analysis 97, S. 103759. DOI: 10.1016/j.irfa.2024.103759.

Burke, Marshall; Hsiang, Solomon M.; Miguel, Edward (2015): Global non-linear effect of temperature on economic production. In: Nature 527 (7577), S. 235–239. DOI: 10.1038/nature15725.

Cai, Yongyang; Lenton, Timothy M.; Lontzek, Thomas S. (2016): Risk of multiple interacting tipping points should encourage rapid CO2 emission reduction. In: Nature Climate Change 6 (5), S. 520–525. DOI: 10.1038/nclimate2964.

Cameron, A. Colin; Gelbach, Jonah B.; Miller, Douglas L. (2011): Robust Inference With Multiway Clustering. In: Journal of Business & Economic Statistics 29 (2), S. 238–249. DOI: 10.1198/jbes.2010.07136.

Campiglio, Emanuele; Daumas, Louis; Monnin, Pierre; Jagow, Adrian von (2023): Climate-related risks in financial assets. In: Journal of Economic Surveys 37 (3), S. 950–992. DOI: 10.1111/joes.12525.

Campopiano, Giovanna; Gabaldón, Patricia; Gimenez-Jimenez, Daniela (2023): Women Directors and Corporate Social Performance: An Integrative Review of the Literature and a Future Research Agenda. In: Journal of Business Ethics 182 (3), S. 717–746. DOI: 10.1007/s10551-021-04999-7.

Carey, Peter; Khan, Arifur; Mihret, Dessalegn G.; Muttakin, Mohammad B. (2021): Voluntary sustainability assurance, capital constraint and cost of debt: International evidence. In: International Journal of Auditing 25 (2), S. 351–372. DOI: 10.1111/ijau.12223.

Carroll, Archie B. (1979): A Three-Dimensional Conceptual Model of Corporate Performance. In: The Academy of Management Review 4 (4), S. 497–505. DOI: 10.5465/amr.1979.4498296.

Carroll, Archie B. (1991): The pyramid of corporate social responsibility: Toward the moral management of organizational stakeholders. In: Business Horizons 34 (4), S. 39–48. DOI: 10.1016/0007-6813(91)90005-G.

Casey, Ryan J.; Grenier, Jonathan H. (2015): Understanding and Contributing to the Enigma of Corporate Social Responsibility (CSR) Assurance in the United States. In: AUDITING: A Journal of Practice & Theory 34 (1), S. 97–130. DOI: 10.2308/ajpt-50736.

Cho, Charles H.; Laine, Matias; Roberts, Robin W.; Rodrigue, Michelle (2015): Organized hypocrisy, organizational façades, and sustainability reporting. In: Accounting, Organizations and Society 40, S. 78–94. DOI: 10.1016/j.aos.2014.12.003.

Christensen, Dane M.; Serafeim, George; Sikochi, Anywhere (2022): Why is Corporate Virtue in the Eye of The Beholder? The Case of ESG Ratings. In: The Accounting Review 97 (1), S. 147–175. DOI: 10.2308/TAR-2019-0506.

Christensen, Hans B.; Hail, Luzi; Leuz, Christian (2021): Mandatory CSR and sustainability reporting: economic analysis and literature review. In: Review of Accounting Studies 26 (3), S. 1176–1248. DOI: 10.1007/s11142-021-09609-5.

Chrzan, Sandra; Pott, Christiane (2024): Limiting environmental reporting flexibility: investor judgment based on the EU taxonomy. In: Review of Quantitative Finance and Accounting 63 (4), S. 1511–1548. DOI: 10.1007/s11156-024-01297-x.

Cohen, Shira; Kadach, Igor; Ormazabal, Gaizka; Reichelstein, Stefan (2023): Executive Compensation Tied to ESG Performance: International Evidence. In: Journal of Accounting Research 61 (3), S. 805–853. DOI: 10.1111/1475-679x.12481.

Courtis, John K. (1995): Readability of annual reports: Western versus Asian evidence. In: Accounting, Auditing & Accountability Journal 8 (2), S. 4–17. DOI: 10.1108/09513579510086795.

Courtis, John K. (1998): Annual report readability variability: tests of the obfuscation hypothesis. In: Accounting, Auditing & Accountability Journal 11 (4), S. 459–472. DOI: 10.1108/09513579810231457.

Courtis, John K. (2004): Corporate report obfuscation: artefact or phenomenon? In: The British Accounting Review 36 (3), S. 291–312. DOI: 10.1016/j.bar.2004.03.005.

Cuadrado-Ballesteros, Beatriz; Martínez-Ferrero, Jennifer; García-Sánchez, Isabel María (2017): Mitigating information asymmetry through sustainability assurance: The role of accountants and levels of assurance. In: International Business Review 26 (6), S. 1141–1156. DOI: 10.1016/j.ibusrev.2017.04.009.

Cuomo, Francesca; Gaia, Silvia; Girardone, Claudia; Piserà, Stefano (2022): The effects of the EU non-financial reporting directive on corporate social responsibility. In: The European Journal of Finance 30 (7), S. 726–752. DOI: 10.1080/1351847X.2022.2113812.

De Franco, Gus; Kothari, S. P.; Verdi, Rodrigo S. (2011): The Benefits of Financial Statement Comparability. In: Journal of Accounting Research 49 (4), S. 895–931. DOI: 10.1111/j.1475-679X.2011.00415.x.

Del Gesso, Carla; Lodhi, Rab Nawaz (2025): Theories underlying environmental, social and governance (ESG) disclosure: a systematic review of accounting studies. In: Journal of Accounting Literature 47 (2), S. 433–461. DOI: 10.1108/JAL-08-2023-0143.

Dell'Erba, Marco; Ferrarini, Guido (2024): ESG & Executive Remuneration in Europe. In: European Business Organization Law Review 25 (3), S. 439–479. DOI: 10.1007/s40804-024-00318-5.

DiMaggio, Paul J.; Powell, Walter W. (1983): The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields. In: American Sociological Review 48 (2), S. 147. DOI: 10.2307/2095101.

Eccles, Robert G.; Krzus, Michael P.; Rogers, Jean; Serafeim, George (2012): The Need for Sector-Specific Materiality and Sustainability Reporting Standards. In: Journal of Applied Corporate Finance 24 (2), S. 65–71. DOI: 10.1111/j.1745-6622.2012.00380.x.

Edmans, Alex; Gabaix, Xavier (2016): Executive Compensation: A Modern Primer. In: Journal of Economic Literature 54 (4), S. 1232–1287. DOI: 10.1257/jel.20161153.

Elkington, John (1998): Partnerships from cannibals with forks: The triple bottom line of 21st-century business. In: Environmental Quality Management 8 (1), S. 37–51. DOI: 10.1002/tqem.3310080106.

Elkington, John (2006): Governance for Sustainability*. In: Corporate Governance 14 (6), S. 522–529. DOI: 10.1111/j.1467-8683.2006.00527.x.

Europäische Kommission (2018): Aktionsplan: Finanzierung nachhaltigen Wachstums. Mitteilung der Kommission an das Europäische Parlament, den Europäischen Rat, den Rat, die Europäische Zentralbank, den Europäischen Wirtschafts- und Sozialausschuss und den Ausschuss der Regionen, COM(2018) 97 final. Brüssel.

Europäische Kommission (2019): Der europäische Grüne Deal. Mitteilung der Kommission an das Europäische Parlament, den Europäischen Rat, den Rat, den Europäischen Wirtschafts- und Sozialausschuss und den Ausschuss der Regionen, COM(2019) 640 final. Brüssel.

Fischer, Erich M.; Knutti, Reto (2015): Anthropogenic contribution to global occurrence of heavy-precipitation and high-temperature extremes. In: Nature Climate Change 5 (6), S. 560–564. DOI: 10.1038/nclimate2617.

Flammer, Caroline; Hong, Bryan; Minor, Dylan (2019): Corporate governance and the rise of integrating corporate social responsibility criteria in executive compensation: Effectiveness and implications for firm outcomes. In: Strategic Management Journal 40 (7), S. 1097–1122. DOI: 10.1002/smj.3018.

Flesch, Rudolf (1948): A New Readability Yardstick. In: Journal of Applied Psychology 32 (3), S. 221–233.

Forbes, Daniel P.; Milliken, Frances J. (1999): Cognition and Corporate Governance: Understanding Boards of Directors as Strategic Decision-Making Groups. In: The Academy of Management Review 24 (3), S. 489–505. DOI: 10.5465/amr.1999.2202133.

Freeman, R. Edward (1984): Strategic Management. A Stakeholder Approach. Cambridge: Cambridge University Press.

Gabaldon, Patricia; Kanadlı, Sadi Boğaç; Bankewitz, Max (2018): How does job-related diversity affect boards' strategic participation? An information-processing approach. In: Long Range Planning 51 (6), S. 937–952. DOI: 10.1016/j.lrp.2018.02.005.

Gallani, Susanna; Krishnan, Ranjani; Wooldridge, Jeffrey M. (2015): Applications of Fractional Response Model to the Study of Bounded Dependent Variables in Accounting Research. Working Paper No. 16-016. Boston: Harvard Business School.

García-Meca, Emma; Ruiz-Barbadillo, Emiliano; Martínez-Ferrero, Jennifer (2026): High-quality assurance, ESG legitimacy threats and board effectiveness. In: The British Accounting Review 58 (4), Art. 101385. DOI: 10.1016/j.bar.2024.101385.

García-Sánchez, Isabel-María; Hussain, Nazim; Aibar-Guzmán, Cristina; Aibar-Guzmán, Beatriz (2021): Assurance of corporate social responsibility reports: Does it reduce decoupling practices? In: Business Ethics, the Environment & Responsibility 31 (1), S. 118–138. DOI: 10.1111/beer.12394.

García-Sánchez, Isabel-María; Hussain, Nazim; Martínez-Ferrero, Jennifer; Ruiz-Barbadillo, Emiliano (2019): Impact of disclosure and assurance quality of corporate sustainability reports on access to finance. In: Corporate Social Responsibility and Environmental Management 26 (4), S. 832–848. DOI: 10.1002/csr.1724.

García-Sánchez, Isabel-María; Marín-Hernández, Salvador; Ortiz-Martínez, Esther; Aibar-Guzmán, Diego (2024): Diversity, Equity, and Inclusion Reporting in European Union Companies: The Role of Female Directors and the European Regulatory Framework. In: Business Strategy and the Environment 33(7): S. 7021–40. DOI: 10.1002/bse.3854.

Garst, Jilde; Maas, Karen; Suijs, Jeroen (2022): Materiality Assessment Is an Art, Not a Science: Selecting ESG Topics for Sustainability Reports. In: California Management Review 65 (1), S. 64–90. DOI: 10.1177/00081256221120692.

Gillan, Stuart L.; Koch, Andrew; Starks, Laura T. (2021): Firms and social responsibility: A review of ESG and CSR research in corporate finance. In: Journal of Corporate Finance 66, S. 101889. DOI: 10.1016/j.jcorpfin.2021.101889.

Gosselin, Anne-Marie; Le Maux, Julien; Smaili, Nadia (2021): Readability of Accounting Disclosures: A Comprehensive Review and Research Agenda*. In: Accounting Perspectives 20 (4), S. 543–581. DOI: 10.1111/1911-3838.12275.

Grewal, Jody; Richardson, Gordon D.; Wang, Jingjing (2026): Effects of Mandatory Carbon Reporting on Greenwashing. In: The Accounting Review 101 (4), S. 263–293. DOI: 10.2308/TAR-2023-0247.

Gunning, Robert (1952): The Technique of Clear Writing. New York: McGraw-Hill.

Hadro, Dominika; Calandra, Davide; Fijalkowska, Justyna; Oppioli, Michele; Secinaro, Silvana (2026): Addressing Symbolic Versus Substantive Disclosures Under CSRD/ESRS E5 in the Circular Economy Disclosure of the Automotive Industry. In: Business Strategy and the Environment 35 (3), S. 4360–4375. DOI: 10.1002/bse.70412.

Hagemeier, Lea; Müller, Maximilian A. (2026): Demand for sustainability disclosures: evidence from web-tracking data of CSRD reports. In: Accounting and Business Research 56 (3), S. 365–397. DOI: 10.1080/00014788.2026.2631933.

Hahn, Rüdiger; Kühnen, Michael (2013): Determinants of sustainability reporting: a review of results, trends, theory, and opportunities in an expanding field of research. In: Journal of Cleaner Production 59, S. 5–21. DOI: 10.1016/j.jclepro.2013.07.005.

Haque, Faizul; Ntim, Collins G. (2020): Executive Compensation, Sustainable Compensation Policy, Carbon Performance and Market Value. In: British Journal of Management 31 (3), S. 525–546. DOI: 10.1111/1467-8551.12395.

Harasheh, Murad; Provasi, Roberta (2023): A need for assurance: Do internal control systems integrate environmental, social, and governance factors? In: Corporate Social Responsibility and Environmental Management 30 (1), S. 384–401. DOI: 10.1002/csr.2361.

Hartmann, Maren; Weißenberger, Barbara E. (2024): Information overload research in accounting: a systematic review of the literature. In: Management Review Quarterly 74 (3), S. 1619–1667. DOI: 10.1007/s11301-023-00343-7.

Hazaea, Saddam A.; Zhu, Jinyu; Khatib, Saleh F. A.; Bazhair, Ayman Hassan; Elamer, Ahmed A. (2022): Sustainability assurance practices: a systematic review and future research agenda. In: Environmental Science and Pollution Research 29 (4), S. 4843–4864. DOI: 10.1007/s11356-021-17359-9.

Healy, Paul M.; Palepu, Krishna G. (2001): Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. In: Journal of

Accounting and Economics 31 (1-3), S. 405–440. DOI: 10.1016/S0165-4101(01)00018-0.

Hemmings, Danial; Hodgkinson, Lynn; Williams, Gwion (2020): It's OK to pay well, if you write well: The effects of remuneration disclosure readability. In: Journal of Business Finance & Accounting 47 (5-6), S. 547–586. DOI: 10.1111/jbfa.12431.

Henkel, Knut; Klokkers, Jann; Gastell, Elly Sophie; Prochno, Sina; Grottel, Gregor (2026): Doppelte Wesentlichkeitsanalyse nach ESRS. Quantifizierung von Schwellenwerten am Beispiel der DAX- und MDAX-Nachhaltigkeitsberichterstattung 2024. Emden: Hochschule Emden/Leer, Fachbereich Wirtschaft. DOI: 10.60771/opus-1149.

Hillman, Amy J.; Dalziel, Thomas (2003): Boards of Directors and Firm Performance: Integrating Agency and Resource Dependence Perspectives. In: The Academy of Management Review 28 (3), S. 383. DOI: 10.2307/30040728.

Hooghiemstra, Reggy; Kuang, Yu Flora; Qin, Bo (2017): Does obfuscating excessive CEO pay work? The influence of remuneration report readability on say-on-pay votes. In: Accounting and Business Research 47 (6), S. 695–729. DOI: 10.1080/00014788.2017.1300516.

Hristov, Ivo; Searcy, Cory (2025): Integrating sustainability with corporate governance: a framework to implement the corporate sustainability reporting directive through a balanced scorecard. In: Management Decision 63 (2), S. 443–467. DOI: 10.1108/md-10-2023-1995.

Hsiang, Solomon; Kopp, Robert; Jina, Amir; Rising, James; Delgado, Michael; Mohan, Shashank et al. (2017): Estimating economic damage from climate change in the United States. In: Science 356 (6345), S. 1362–1369. DOI: 10.1126/science.aal4369.

Hummel, Katrin; Jobst, Dominik (2024): An Overview of Corporate Sustainability Reporting Legislation in the European Union. In: Accounting in Europe 21 (3), S. 320–355. DOI: 10.1080/17449480.2024.2312145.

Hussain, Nazim; Rigoni, Ugo; Orij, René P. (2018): Corporate Governance and Sustainability Performance: Analysis of Triple Bottom Line Performance. In: Journal of Business Ethics 149 (2), S. 411–432. DOI: 10.1007/s10551-016-3099-5.

IDW (2018): Stellungnahme des Instituts der Wirtschaftsprüfer zum Entwurf eines Gesetzes zur Umsetzung der zweiten Aktionärsrechterichtlinie (ARUG II)

bmj.de/SharedDocs/Downloads/DE/Gesetzgebung/Stellungnahmen/2019/12042018_Stellungnahme_Institut-Wirtschaftspruefer_ARUG-II.pdf?__blob=publicationFile (abgerufen am 07.09. 2026).

Ilhan, Emirhan; Sautner, Zacharias; Vilkov, Grigory (2021): Carbon Tail Risk. In: The Review of Financial Studies 34 (3), S. 1540–1571. DOI: 10.1093/rfs/hhaa071.

Ioannou, Ioannis; Serafeim, George (2017): The Consequences of Mandatory Corporate Sustainability Reporting. DOI: 10.2139/SSRN.1799589.

IPCC (2023): Climate Change 2023 Synthesis Report. Online verfügbar unter https://www.ipcc.ch/report/ar6/syr/.

Jensen, Michael C.; Meckling, William H. (1976): Theory of the firm: Managerial behavior, agency costs and ownership structure. In: Journal of Financial Economics 3 (4), S. 305–360. DOI: 10.1016/0304-405X(76)90026-X.

Kanadlı, Sadi Boğaç; Bankewitz, Max; Zhang, Pingying (2018): Job-related diversity: the comprehensiveness and speed of board decision-making processes—an upper echelons approach. In: Journal of Management and Governance 22 (2), S. 427–456. DOI: 10.1007/s10997-017-9394-4.

Kanter, Rosabeth Moss (1977): Some Effects of Proportions on Group Life: Skewed Sex Ratios and Responses to Token Women. In: American Journal of Sociology 82 (5), S. 965–990. DOI: 10.1086/226425.

Khanna, Poonam; Jones, Carla D.; Boivie, Steven (2014): Director Human Capital, Information Processing Demands, and Board Effectiveness. In: Journal of Management 40 (2), S. 557–585. DOI: 10.1177/0149206313515523.

Kim, Hajin; Li, Ningzi; Feldman, Ronen; Liu, Yun; Feldman, Yuval (2026): What Sustainability Disclosures Disclose. Coase-Sandor Institute for Law & Economics Research Paper No. 26-7. Chicago: University of Chicago Law School. Online verfügbar unter https://ssrn.com/abstract=7111640.

Kölbel, Julian F.; Heeb, Florian; Paetzold, Falko; Busch, Timo (2020): Can Sustainable Investing Save the World? Reviewing the Mechanisms of Investor Impact. In: Organization & Environment 33 (4), S. 554–574. DOI: 10.1177/1086026620919202.

Kolk, Ans (2008): Sustainability, accountability and corporate governance: exploring multinationals' reporting practices. In: Business Strategy and the Environment 17 (1), S. 1–15. DOI: 10.1002/bse.511.

Kolk, Ans; Perego, Paolo (2010): Determinants of the adoption of sustainability assurance statements: an international investigation. In: Business Strategy and the Environment 19 (3), S. 182–198. DOI: 10.1002/bse.643.

Korca, Blerita; Costa, Ericka; Bouten, Lies (2023): Disentangling the concept of comparability in sustainability reporting. In: Sustainability Accounting, Management and Policy Journal 14 (4), S. 815–851. DOI: 10.1108/sampj-05-2022-0284.

KPMG (2022): Big shifts, small steps. Survey of Sustainability Reporting 2022. Online verfügbar unter https://kpmg.com/sk/en/insights/2022/09/survey-of-sustainability-reporting-2022.html.

Krueger, Philipp; Sautner, Zacharias; Starks, Laura T. (2020): The Importance of Climate Risks for Institutional Investors. In: The Review of Financial Studies 33 (3), S. 1067–1111. DOI: 10.1093/rfs/hhz137.

Krueger, Philipp; Sautner, Zacharias; Tang, Dragon Yongjun; Zhong, RuI. (2024): The Effects of Mandatory ESG Disclosure Around the World. In: Journal of Accounting Research 62 (5), S. 1795–1847. DOI: 10.1111/1475-679X.12548.

Kuckartz, Udo; Rädiker, Stefan (2024): Fokussierte Interviewanalyse mit MAXQDA. Wiesbaden: Springer Fachmedien Wiesbaden.

Kuthe, Thorsten; Reuters, Madeleine (2022): Der aktienrechtliche Vergütungsbericht nach § 162 AktG. Praxisrelevante Hinweise. In: Betriebs-Berater 77 (37), S. 2115–2120.

La Torre, Matteo; Sabelfeld, Svetlana; Blomkvist, Marita; Tarquinio, Lara; Dumay, John (2018): Harmonising non-financial reporting regulation in Europe. In: Meditari Accountancy Research 26 (4), S. 598–621. DOI: 10.1108/medar-02-2018-0290.

Laksmana, Indrarini; Tietz, Wendy; Yang, Ya-Wen (2012): Compensation discussion and analysis (CD&A): Readability and management obfuscation. In: Journal of Accounting and Public Policy 31 (2), S. 185–203. DOI: 10.1016/j.jaccpubpol.2011.08.003.

Lambert, Richard; Leuz, Christian; Verrecchia, Robert E. (2007): Accounting Information, Disclosure, and the Cost of Capital. In: Journal of Accounting Research 45 (2), S. 385–420. DOI: 10.1111/j.1475-679x.2007.00238.x.

Laub, Johannes; Steller, Marcel (2026): Navigating regulatory ambiguity: evolving challenges and the sustainability expectation gap in assurance under the corporate sustainability reporting directive. In: Sustainability Accounting, Management and Policy Journal, S. 1–29. DOI: 10.1108/sampj-05-2025-0652.

Lenton, Timothy M.; Held, Hermann; Kriegler, Elmar; Hall, Jim W.; Lucht, Wolfgang; Rahmstorf, Stefan; Schellnhuber, Hans Joachim (2008): Tipping elements in the Earth's climate system. In: Proceedings of the National Academy of Sciences of the United States of America 105 (6), S. 1786–1793. DOI: 10.1073/pnas.0705414105.

Lenton, Timothy M.; Rockström, Johan; Gaffney, Owen; Rahmstorf, Stefan; Richardson, Katherine; Steffen, Will; Schellnhuber, Hans Joachim (2019): Climate tipping points - too risky to bet against. In: Nature 575 (7784), S. 592–595. DOI: 10.1038/d41586-019-03595-0.

Li, Feng (2008): Annual report readability, current earnings, and earnings persistence. In: Journal of Accounting and Economics 45 (2-3), S. 221–247. DOI: 10.1016/j.jacceco.2008.02.003.

Lo, Kin & Ramos, Felipe & Rogo, Rafael, 2017: Earnings management and annual report readability, In: Journal of Accounting and Economics 63(1), S.1-25. DOI: 10.1016/j.jacceco.2016.09.002.

Lu, Longjie (2023): ESG-based remuneration in the wave of sustainability. In: Journal of Corporate Law Studies 23 (1), S. 297–339. DOI: 10.1080/14735970.2023.2253888.

Ludwig, Paul; Sassen, Remmer (2022): Which internal corporate governance mechanisms drive corporate sustainability? In: Journal of Environmental Management 301, S. 113780. DOI: 10.1016/j.jenvman.2021.113780.

Marín-Hernández, Salvador; Fernández-Martínez, Pascual; Ortiz-Martínez, Esther (2026): Double Materiality in European Water-Sector Companies: Evidence from the First Application of the European Sustainability Reporting Standards. In: World 7 (7), S. 106. DOI: 10.3390/world7070106.

Marshall, Andrew P.; Weetman, Pauline (2002): Information asymmetry in disclosure of foreign exchange risk management: can regulation be effective? In: Journal of Economics and Business 54 (1), S. 31–53. DOI: 10.1016/s0148-6195(01)00058-3.

Martínez-Ferrero, Jennifer; García-Sánchez, Isabel-María (2017): Coercive, normative and mimetic isomorphism as determinants of the voluntary assurance of sustainability reports. In: International Business Review 26 (1), S. 102–118. DOI: 10.1016/j.ibusrev.2016.05.009.

Martínez-Ferrero, Jennifer; García-Sánchez, Isabel-María; Ruiz-Barbadillo, Emiliano (2018): The quality of sustainability assurance reports: The expertise and experience of assurance providers as determinants. In: Business Strategy and the Environment 27 (8), S. 1181–1196. DOI: 10.1002/bse.2061.

Masi, Sara de; Słomka-Gołębiowska, Agnieszka; Becagli, Claudio; Paci, Andrea (2021): Toward sustainable corporate behavior: The effect of the critical mass of female directors on environmental, social, and governance disclosure. In: Business Strategy and the Environment 30 (4), S. 1865–1878. DOI: 10.1002/bse.2721.

Mayring, Philipp (2023): Einführung in die qualitative Sozialforschung. Eine Anleitung zu qualitativem Denken. 7., überarbeitete Auflage. Weinheim, Basel: Beltz. Online verfügbar unter https://www.content-select.com/index.php?id=bib_view&ean=9783407296023.

Mercer, Molly (2004): How Do Investors Assess the Credibility of Management Disclosures? In: Accounting Horizons 18 (3), S. 185–196. DOI: 10.2308/acch.2004.18.3.185.

Meyer, John W.; Rowan, Brian (1977): Institutionalized Organizations: Formal Structure as Myth and Ceremony. In: American Journal of Sociology 83 (2), S. 340–363. DOI: 10.1086/226550.

Michelon, Giovanna; Parbonetti, Antonio (2012): The effect of corporate governance on sustainability disclosure. In: Journal of Management and Governance 16 (3), S. 477–509. DOI: 10.1007/s10997-010-9160-3.

Miettinen, Mirella (2024): Are materiality determination practices evolving in the wake of increasing legislation on sustainability reporting? Findings from EU pharmaceutical companies' reports. In: International Journal of Law and Management 66 (3), S. 363–392. DOI: 10.1108/ijlma-09-2023-0221.

Min, Seung-Ki; Zhang, Xuebin; Zwiers, Francis W.; Hegerl, Gabriele C. (2011): Human contribution to more-intense precipitation extremes. In: Nature 470 (7334), S. 378–381. DOI: 10.1038/nature09763.

Morganho, Catarina; Milánes-Montero, Patricia; Pérez-Calderón, Esteban (2025): The Determinants of Materiality Assessment Disclosures in the Sustainability Statements of Iberian-Listed Companies. In: Corporate Social Responsibility and Environmental Management 32 (3), S. 3659–3673. DOI: 10.1002/csr.3148.

Naciti, Valeria; Cesaroni, Fabrizio; Pulejo, Luisa (2021): Corporate governance and sustainability: a review of the existing literature. In: Journal of Management and Governance 26 (1), S. 55–74. DOI: 10.1007/s10997-020-09554-6.

Nadeem, Muhammad (2022): Board Gender Diversity and Managerial Obfuscation: Evidence from the Readability of Narrative Disclosure in 10-K Reports. In: Journal of Business Ethics 179 (1), S. 153–177. DOI: 10.1007/s10551-021-04830-3.

Nahrstedt, Jan (2026): Double Materiality in Practice: Evidence from Two Years of CSRD Reporting (May 06, 2026). Online verfügbar unter https://ssrn.com/abstract=6724699. DOI: 10.2139/ssrn.6724699.

Nicolò, Giuseppe; Zampone, Giovanni; Sannino, Giuseppe; Iorio, Serena de (2022): Sustainable corporate governance and non-financial disclosure in Europe: does the gender diversity matter? In: Journal of Applied Accounting Research 23 (1), S. 227–249. DOI: 10.1108/JAAR-04-2021-0100.

Nipper, Marvin; Ostermaier, Andreas; Theis, Jochen (2025): Mandatory disclosure of standardized sustainability metrics: The case of the EU Taxonomy Regulation. In: Corporate Social Responsibility and Environmental Management 32 (2), S. 2171–2190. DOI: 10.1002/csr.3046.

OECD (2023): G20/OECD Principles of Corporate Governance 2023, Paris: OECD Publishing. DOI: 10.1787/ed750b30-en.

Ottenstein, Philipp; Erben, Saskia; Jost, Sébastien; Weuster, Carl William; Zülch, Henning (2022): From voluntarism to regulation: effects of Directive 2014/95/EU on sustainability reporting in the EU. In: Journal of Applied Accounting Research 23 (1), S. 55–98. DOI: 10.1108/jaar-03-2021-0075.

Pacces, Alessio M. (2021): Will the EU Taxonomy Regulation Foster Sustainable Corporate Governance? In: Sustainability 13 (21), S. 12316. DOI: 10.3390/su132112316.

Pacelli, Vincenzo; Ekinci, Selen (2026): ESRS Implementation under the CSRD : A Comparative Directed Qualitative Content Analysis of Ten Corporate Cases. In:

Corporate Social Responsibility and Environmental Management, Artikel csr.70813. DOI: 10.1002/csr.70813.

Papke, Leslie E.; Wooldridge, Jeffrey M. (1996): Econometric methods for fractional response variables with an application to 401(k) plan participation rates. In: Journal of Applied Econometrics 11 (6), S. 619–632. DOI: 10.1002/(SICI)1099-1255(199611)11:6<619::AID-JAE418>3.0.CO;2-1.

Perego, Paolo; Kolk, Ans (2012): Multinationals' Accountability on Sustainability: The Evolution of Third-party Assurance of Sustainability Reports. In: Journal of Business Ethics 110 (2), S. 173–190. DOI: 10.1007/s10551-012-1420-5.

Petersen, Mitchell A. (2009): Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches. In: The Review of Financial Studies 22 (1), S. 435–480. DOI: 10.1093/rfs/hhn053.

Pfeffer, Jeffrey; Salancik, Gerald R. (1978): The External Control of Organizations. A Resource Dependence Perspective. New York: Harper & Row.

Pirson, Michael; Turnbull, Shann (2011): Corporate Governance, Risk Management, and the Financial Crisis: An Information Processing View. In: Corporate Governance 19 (5), S. 459–470. DOI: 10.1111/j.1467-8683.2011.00860.x.

Quick, Reiner (2024): Zur Ausgestaltung der Pflicht zur Prüfung von Nachhaltigkeitsberichten. Anmerkungen zu den im Referentenentwurf des CSRD-UmsG vorgesehenen Regelungen zur Pflichtprüfung. In: WPg – Die Wirtschaftsprüfung (9), S. 461–516.

Quick, Reiner; Sánchez Toledano, Daniel; Sánchez Toledano, Joaquín (2022): Ist eine Prüfungspflicht für Nachhaltigkeitsberichte sinnvoll? Eine wissenschaftsbasierte Diskussion des EU-Vorschlags zu einer überarbeiteten CSR-Richtlinie. In: IRZ – Zeitschrift für Internationale Rechnungslegung (1), S. 3–42.

Raith, Dirk (2023): The contest for materiality. What counts as CSR? In: Journal of Applied Accounting Research 24 (1), S. 134–148. DOI: 10.1108/jaar-04-2022-0093.

Reimsbach, Daniel; Hahn, Rüdiger; Gürtürk, Anil (2018): Integrated Reporting and Assurance of Sustainability Information: An Experimental Study on Professional Investors' Information Processing. In: European Accounting Review 27 (3), S. 559–581. DOI: 10.1080/09638180.2016.1273787.

Rennekamp, Kristina (2012): Processing Fluency and Investors' Reactions to Disclosure Readability. In: Journal of Accounting Research 50 (5), S. 1319–1354. DOI: 10.1111/j.1475-679X.2012.00460.x.

Rennert, Kevin; Errickson, Frank; Prest, Brian C.; Rennels, Lisa; Newell, Richard G.; Pizer, William et al. (2022): Comprehensive evidence implies a higher social cost of CO2. In: Nature 610 (7933), S. 687–692. DOI: 10.1038/s41586-022-05224-9.

Richardson, Katherine; Steffen, Will; Lucht, Wolfgang; Bendtsen, Jørgen; Cornell, Sarah E.; Donges, Jonathan F. et al. (2023): Earth beyond six of nine planetary boundaries. In: Science Advances 9 (37), eadh2458. DOI: 10.1126/sciadv.adh2458.

Rockström, Johan; Gupta, Joyeeta; Qin, Dahe; Lade, Steven J.; Abrams, Jesse F.; Andersen, Lauren S. et al. (2023): Safe and just Earth system boundaries. In: Nature 619 (7968), S. 102–111. DOI: 10.1038/s41586-023-06083-8.

Rouf, Md. Abdur; Siddique, Md. Nur-E-Alam (2023): Theories applied in corporate voluntary disclosure: a literature review. In: Journal of Entrepreneurship and Public Policy 12 (1), S. 49–68. DOI: 10.1108/jepp-01-2022-0007.

Ruiz-Barbadillo, Emiliano; Martínez-Ferrero, Jennifer (2020): Empirical analysis of the effect of the joint provision of audit and sustainability assurance services on assurance quality. In: Journal of Cleaner Production 266, S. 121943. DOI: 10.1016/j.jclepro.2020.121943.

Ruohonen, Janne; Kullas, Helmi (2024): 442The Assurance of Corporate Sustainability Reports and the Renewed Role of Certified Auditors. In: European Company and Financial Law Review 21 (3-4), S. 442–471. DOI: 10.1515/ecfr-2024-0013.

Ruohonen, Janne; Tamlander, Eero (2025): EU corporate sustainability legislation – from expansion to Omnibus. In: ECCL 3 (4), S. 381–402. DOI: 10.5771/2752-177x-2025-4-381.

Sabbatucci, Agnese; Guidi, Michele; Giuliani, Marco (2026): Sustainability assurance landscape after the CSRD: scenario analysis and evolutionary pathways in Europe. In: Sustainability Accounting, Management and Policy Journal 17 (3), S. 984–1011. DOI: 10.1108/sampj-12-2024-1393.

Schick, Allen G.; Gordon, Lawrence A.; Haka, Susan (1990): Information overload: A temporal approach. In: Accounting, Organizations and Society 15 (3), S. 199–220. DOI: 10.1016/0361-3682(90)90005-F.

Schiehll, Eduardo; Lewellyn, Krista; Yan, Wenxi (2023): A configurational perspective of boards' attention structures. In: Corporate Governance 31 (5), S. 676–696. DOI: 10.1111/corg.12493.

Schmidt, Matthias; Schmotz, Thomas; von der Heide, Marten (2025): Die Praxis der Nachhaltigkeitsberichterstattung nach CSRD/ESRS. Eine Analyse von DAX-, MDAX- und SDAX-Unternehmen für das Geschäftsjahr 2024. In: Der Betrieb 78 (29), S. 1705–1712.

Schütze, Franziska; Sandbaek, Benedikte (2025): Taxonomy Disclosure in the EU – A Useful Framework, Despite Current Challenges. In: The Economists' Voice 22 (1), S. 161–172. DOI: 10.1515/ev-2025-0016.

Seele, Peter; Gatti, Lucia (2017): Greenwashing Revisited: In Search of a Typology and Accusation-Based Definition Incorporating Legitimacy Strategies. In: Business Strategy and the Environment 26 (2), S. 239–252. DOI: 10.1002/bse.1912.

Sellhorn, Thorsten; Wagner, Victor (2022): The Forces that Shape Mandatory ESG Reporting (July 30, 2022). TRR 266 Accounting for Transparency Working Paper Series No. 96, 2022, Online verfügbar unter https://ssrn.com/abstract=4179479. DOI: 10.2139/ssrn.4179479.

Seo, Jeongil (2017): Board Effectiveness and CEO Pay: Board Information Processing Capacity, Monitoring Complexity, and CEO Pay-for-Performance Sensitivity. In: Human Resource Management 56 (3), S. 373–388. DOI: 10.1002/hrm.21769.

Shleifer, Andrei; Vishny, Robert W. (1997): A Survey of Corporate Governance. In: The Journal of Finance 52 (2), S. 737–783. DOI: 10.1111/j.1540-6261.1997.tb04820.x.

Simnett, Roger; Vanstraelen, Ann; Chua, Wai Fong (2009): Assurance on Sustainability Reports: An International Comparison. In: The Accounting Review 84 (3), S. 937–967. DOI: 10.2308/accr.2009.84.3.937.

Simoni, Lorenzo; Bini, Laura; Bellucci, Marco (2020): Effects of social, environmental, and institutional factors on sustainability report assurance: evidence from European countries. In: Meditari Accountancy Research 28 (6), S. 1059–1087. DOI: 10.1108/medar-03-2019-0462.

Sopp, Karina; Baumüller, Josef; Scheid, Oliver (2022): Nachhaltigkeitsberichterstattung. Nichtfinanzielle Berichterstattung nach dem CSR-RUG und Neuerungen durch die CSRD. 2. Aufl. Herne: NWB.

Spallek, Franziska; Bastini, Karola (2026): Linking Executive Pay to ESG Criteria: An Analysis of Current Practice in German Listed Firms. In: Schmalenbach Journal of Business Research 78 (1). DOI: 10.1007/s41471-026-00236-x.

Stanek-Kowalczyk, Aleksandra; Szumniak-Samolej, Justyna (2026): Omnibus I and the Challenges of Sustainability Reporting in the EU. In: Corporate Social Responsibility and Environmental Management, Artikel csr.70610. DOI: 10.1002/csr.70610.

Staubus, George J. (1961): A Theory of Accounting to Investors. Berkeley: University of California Press.

Staubus, George J. (2013): The Decision Usefulness Theory of Accounting. A Limited History. London: Routledge.

Ștefănescu, Cristina Alexandrina (2022): Linking sustainability and non-financial reporting directive 2014/95/EU through isomorphism lens. In: Meditari Accountancy Research 30 (6), S. 1680–1704. DOI: 10.1108/medar-09-2020-1019.

Stellner, Bernhard (2022): Readability of Annual Reports on the Vienna Stock Exchange: A Test of Management Obfuscation Hypothesis. In: CEBR 11 (5), S. 49–66. DOI: 10.18267/j.cebr.307.

Steuer, Sebastian; Tröger, Tobias H. (2022): The Role of Disclosure in Green Finance. In: Journal of Financial Regulation 8 (1), S. 1–50. DOI: 10.1093/jfr/fjac001.

Suchman, Mark C. (1995): Managing Legitimacy: Strategic and Institutional Approaches. In: The Academy of Management Review 20 (3), S. 571–610. DOI: 10.5465/amr.1995.9508080331.

Tamimi, Nabil; Sebastianelli, Rose (2017): Transparency among S&P 500 companies: an analysis of ESG disclosure scores. In: Management Decision 55 (8), S. 1660–1680. DOI: 10.1108/MD-01-2017-0018.

Tan, Hun-Tong; Wang, Elaine Ying; Zhou, Bo (2014): When the Use of Positive Language Backfires: The Joint Effect of Tone, Readability, and Investor Sophistication on Earnings Judgments. In: Journal of Accounting Research 52 (1), S. 273–302. DOI: 10.1111/1475-679x.12039.

Thoms, Claudia; Degenhart, Anke; Wohlgemuth, Katharina (2020): Is Bad News Difficult to Read? A Readability Analysis of Differently Connoted Passages in the Annual Reports of the 30 DAX Companies. In: Journal of Business and Technical Communication 34 (2), S. 157–187. DOI: 10.1177/1050651919892312.

Tilbury, Louise; Sealy, Ruth (2024): Seen and not heard: A comparative case study of women on boards and process loss beyond critical mass. In: Corporate Governance 32 (2), S. 190–204. DOI: 10.1111/corg.12524.

Tol, Richard S. J. (2023): Social cost of carbon estimates have increased over time. In: Nature Climate Change 13 (6), S. 532–536. DOI: 10.1038/s41558-023-01680-x.

Tutino, Marco; Arduini, Simona; Beck, Tommaso; Capolini, Francesca (2025): Sustainable corporate governance: A structured literature review of the past 10 years. In: Corporate Ownership and Control 22 (3), S. 98. DOI: 10.22495/cocv22i3art8.

UNFCCC (2015): Paris Agreement. FCCC/CP/2015/10/Add.1. 21st Conference of the Parties (COP21). Online verfügbar unter https://unfccc.int/process-and-meetings/the-paris-agreement.

United Nations (2015): Transforming our World. The 2030 Agenda for Sustainable Development. Resolution der Generalversammlung vom 25. September 2015, A/RES/70/1.

Vander Bauwhede, Heidi; van Cauwenberge, Philippe (2022): Determinants and Value Relevance of Voluntary Assurance of Sustainability Reports in a Mandatory Reporting Context: Evidence from Europe. In: Sustainability 14 (15), S. 9795. DOI: 10.3390/su14159795.

Velte, Patrick (2021a): Regulierung der nachhaltigen Corporate Governance? Eine kritische Würdigung unter Berücksichtigung des geplanten Finanzmarktintegritätsstärkungsgesetzes (FISG). In: Neue Zeitschrift für Gesellschaftsrecht 24 (1), S. 3–5.

Velte, Patrick (2021b): Effect of overlapping audit and compensation committee memberships on the readability of management compensation reports in the German HDAX. In: Journal of Managerial Issues, 33. Jg., S. 27–48.

Velte, Patrick (2022a): Does sustainable corporate governance have an impact on materiality disclosure quality in integrated reporting? International evidence. In: Sustainable Development 30 (6), S. 1655–1670. DOI: 10.1002/sd.2333.

Velte, Patrick (2022b): Prüfung der Nachhaltigkeitsberichterstattung durch den Abschlussprüfer. Anmerkungen zur geplanten Überarbeitung der CSR-Richtlinie und zum aktuellen Durand Report. In: WPg – Die Wirtschaftsprüfung (2), S. 61–112.

Velte, Patrick (2023): Auswahl des Prüfungsdienstleisters für Nachhaltigkeitsberichte nach der CSRD. In: WPg – Die Wirtschaftsprüfung (23), S. 1285–1348.

Velte, Patrick (2025a): Board Gender Diversity and Corporate Social Responsibility. A Literature Review on Critical Mass Theory and Research Recommendations. In: Sustainable Development 33 (S1), S. 1233–1248. DOI: 10.1002/sd.70071.

Velte, Patrick (2025b): Determinants of the selection of sustainability assurance providers and consequences for firm value: a review of empirical research. In: Meditari Accountancy Research 33 (7), S. 443–465. DOI: 10.1108/medar-07-2024-2575.

Velte, Patrick (2026): Sustainability reporting regulations at the crossroads. A critical note on the EU Omnibus. In: Critical Perspectives on International Business, S. 1–23. DOI: 10.1108/cpoib-06-2025-0142.

Velte, Patrick; Stawinoga, Martin (2022): Prüfung von Nachhaltigkeitsberichten nach der Corporate Sustainability Reporting Directive (CSRD). Zum Nachhaltigkeitsbericht nach den geplanten ESRS aus der Sicht von Aufsichtsrat und Prüfungsausschuss. In: WPg – Die Wirtschaftsprüfung (19).

Venter, Elmar R.; Krasodomska, Joanna (2024): Research on extended external reporting assurance: An update on recent developments. In: Journal of International Financial Management & Accounting 35 (2), S. 390–428. DOI: 10.1111/jifm.12200.

Vetter, Eberhard; Tielmann, Jörgen (2022): Zum Standort des Vergütungsberichts nach § 162 AktG. In: Neue Zeitschrift für Gesellschaftsrecht 25 (8), S. 387–388.

Villiers, Charl de; La Torre, Matteo; Molinari, Matteo (2022): The Global Reporting Initiative's (GRI) past, present and future: critical reflections and a research agenda on sustainability reporting (standard-setting). In: PAR 34 (5), S. 728–747. DOI: 10.1108/par-02-2022-0034.

Vishnu Nampoothiri, Madhavan; Kumar, Mohit (2026): When one size doesn't fit all: how national standards and scope choices fragment CSRD assurance. In: Meditari Accountancy Research, S. 1–20. DOI: 10.1108/medar-11-2025-3409.

Weichbold, Martin (2022): Pretests. In: Nina Baur und Jörg Blasius (Hg.): Handbuch Methoden der empirischen Sozialforschung. Wiesbaden: Springer Fachmedien Wiesbaden, S. 443–451.

Wijen, Frank (2014): Means versus Ends in Opaque Institutional Fields: Trading off Compliance and Achievement in Sustainability Standard Adoption. In: The Academy of Management Review 39 (3), S. 302–323. DOI: 10.5465/amr.2012.0218.

Wooldridge, Jeffrey M. (2010): Econometric Analysis of Cross Section and Panel Data. 2. Aufl. Cambridge, MA: MIT Press.

Wunderling, Nico; Heydt, Anna S. von der; Aksenov, Yevgeny; Barker, Stephen; Bastiaansen, Robbin; Brovkin, Victor et al. (2024): Climate tipping point interactions and cascades: a review. In: Earth System Dynamics 15 (1), S. 41–74. DOI: 10.5194/esd-15-41-2024.

Zamil, Izdihar Abdullah; Ramakrishnan, Suresh; Jamal, Noriza Mohd; Hatif, Majeed Abdulhussein; Khatib, Saleh F.A. (2023): Drivers of corporate voluntary disclosure: a systematic review. In: Journal of Financial Reporting and Accounting 21 (2), S. 232–267. DOI: 10.1108/jfra-04-2021-0110.

Zattoni, Alessandro; Leventis, Stergios; van Ees, Hans; Masi, Sara de (2023): Board diversity's antecedents and consequences: A review and research agenda. In: The Leadership Quarterly 34 (1), S. 101659. DOI: 10.1016/j.leaqua.2022.101659.

Zengin-Karaibrahimoglu, Yasemin; Homroy, Swarnodeep; Lobo, Gerald J. (2026): What do we know about sustainability reporting assurance quality? In: The British Accounting Review 58 (4). DOI: 10.1016/j.bar.2026.101917.

Zhang, Pingying (2010): Board Information and Strategic Tasks Performance. In: Corporate Governance 18 (5), S. 473–487. DOI: 10.1111/j.1467-8683.2010.00816.x.

Zorio, Ana; García-Benau, Maria A.; Sierra, Laura (2013): Sustainability Development and the Quality of Assurance Reports: Empirical Evidence. In: Business Strategy and the Environment 22 (7), S. 484–500. DOI: 10.1002/bse.1764.

Richtlinien, Verordnungen, Gesetze und Standards

Richtlinie 2014/95/EU des Europäischen Parlaments und des Rates vom 22. Oktober 2014 zur Änderung der Richtlinie 2013/34/EU im Hinblick auf die Angabe nichtfinanzieller und die Diversität betreffender Informationen durch bestimmte große Unternehmen und Gruppen (NFRD), ABl. L 2014/330, 15.11.2014.

Richtlinie (EU) 2017/828 des Europäischen Parlaments und des Rates vom 17. Mai 2017 zur Änderung der Richtlinie 2007/36/EG im Hinblick auf die Förderung der langfristigen Mitwirkung der Aktionäre (SRD II), ABl. L 2017/132, 20.05.2017.

Verordnung (EU) 2019/876 des Europäischen Parlaments und des Rates vom 20. Mai 2019 zur Änderung der Verordnung (EU) Nr. 575/2013 hinsichtlich der Verschuldungsquote, der strukturellen Liquiditätsquote, der Anforderungen an Eigenmittel und berücksichtigungsfähige Verbindlichkeiten, des Gegenparteiausfallrisikos, des Marktrisikos, der Risikopositionen gegenüber zentralen Gegenparteien, der Risikopositionen gegenüber Organismen für gemeinsame Anlagen, der Großkredite, der Melde- und Offenlegungspflichten sowie der Verordnung (EU) Nr. 648/2012 (CRR II), ABl. L 2019/150, 07.06.2019.

Verordnung (EU) 2019/2088 des Europäischen Parlaments und des Rates vom 27. November 2019 über nachhaltigkeitsbezogene Offenlegungspflichten im Finanzdienstleistungssektor (SFDR), ABl. L 2019/317, 09.12.2019.

Verordnung (EU) 2020/852 des Europäischen Parlaments und des Rates vom 18. Juni 2020 über die Einrichtung eines Rahmens zur Erleichterung nachhaltiger Investitionen und zur Änderung der Verordnung (EU) 2019/2088 (Taxonomie-Verordnung), ABl. L 2020/198, 22.06.2020.

Delegierte Verordnung (EU) 2021/2178 der Kommission vom 6. Juli 2021 zur Ergänzung der Verordnung (EU) 2020/852 des Europäischen Parlaments und des Rates durch Festlegung des Inhalts und der Darstellung der Informationen, die von Unternehmen, die Artikel 19a oder 29a der Richtlinie 2013/34/EU unterliegen, in Bezug auf ökologisch nachhaltige Wirtschaftstätigkeiten offenzulegen sind, und der Methode, anhand deren die Einhaltung dieser Offenlegungspflicht zu gewährleisten ist, ABl. L 2022/443, 10.12.2021.

Durchführungsverordnung (EU) 2022/2453 der Kommission vom 30. November 2022 zur Änderung der Durchführungsverordnung (EU) 2021/637 hinsichtlich der Offenlegung von ESG-Risiken, ABl. L 2022/324, 19.12.2022.

Richtlinie (EU) 2022/2464 des Europäischen Parlaments und des Rates vom 14. Dezember 2022 zur Änderung der Verordnung (EU) Nr. 537/2014 und der Richtlinien 2004/109/EG, 2006/43/EG und 2013/34/EU hinsichtlich der Nachhaltigkeitsberichterstattung von Unternehmen (CSRD), ABl. L 2022/322, 16.12.2022.

Delegierte Verordnung (EU) 2023/2772 der Kommission vom 31. Juli 2023 zur Ergänzung der Richtlinie 2013/34/EU des Europäischen Parlaments und des Rates hinsichtlich Standards für die Nachhaltigkeitsberichterstattung (ESRS Set 1), ABl. L, 2023/2772, 22.12.2023.

Delegierte Verordnung (EU) 2025/1416 der Kommission vom 11. Juli 2025 zur Änderung der Delegierten Verordnung (EU) 2023/2772 im Hinblick auf die Verlängerung von Übergangserleichterungen (ESRS-Übergangserleichterungen), ABl. L, 2025/1416.

Entwurf eines Gesetzes zur Umsetzung der Richtlinie (EU) 2022/2464 hinsichtlich der Nachhaltigkeitsberichterstattung von Unternehmen in der durch die Richtlinie (EU) 2025/794 geänderten Fassung (CSRD-Umsetzungsgesetz), Regierungsentwurf vom 3. September 2025, BT-Drs. 21/1857.

Richtlinie (EU) 2026/470 des Europäischen Parlaments und des Rates vom TT. Monat 2026 zur Änderung der Richtlinien 2006/43/EG und 2013/34/EU hinsichtlich der Anforderungen an die Nachhaltigkeitsberichterstattung (Omnibus I), ABl. L, 2026/470, 26.02.2026.

Delegierte Verordnung der Kommission vom 3. Juli 2026 zur Änderung der Delegierten Verordnung (EU) 2023/2772 hinsichtlich der überarbeiteten Standards für die Nachhaltigkeitsberichterstattung (ESRS 2026), C(2026) 5010 final.

International Auditing and Assurance Standards Board (2013): International Standard on Assurance Engagements (ISAE) 3000 (Revised), Assurance Engagements Other than Audits or Reviews of Historical Financial Information.

International Auditing and Assurance Standards Board (2024): International Standard on Sustainability Assurance (ISSA) 5000, General Requirements for Sustainability Assurance Engagements. Online verfügbar unter https://www.iaasb.org/focus-areas/understanding-international-standard-sustainability-assurance-5000.

2. Die Prüfung der doppelten Wesentlichkeitsanalyse im Kontext der Nachhaltigkeitsberichterstattung

2.1 Publikationsdetails

Autoren: Jonas Maiwald, Ilka Rogowski

Zusammenfassung: Angesichts des wachsenden Bedarfs an zuverlässigen Nachhaltigkeitsinformationen führt die Corporate Sustainability Reporting Directive (CSRD) erweiterte Berichtspflichten sowie eine externe Prüfungspflicht für einen deutlich größeren Anwenderkreis ein. Dabei spielt das überarbeitete Konzept der doppelten Wesentlichkeitsanalyse eine zentrale Rolle, um die relevanten Berichtsinhalte festzulegen. Basierend auf 12 Experteninterviews leistet diese Arbeit einen Beitrag zum besseren Verständnis der erforderlichen Prüfungsprozesse. Die Prüfung der doppelten Wesentlichkeitsanalyse erfordert eine intensive Auseinandersetzung mit branchenspezifischen Risiken und eine umfassende Analyse der zugrunde liegenden Daten. Während die Prüfung mit begrenzter Sicherheit vorwiegend auf Plausibilitätsannahmen und analytische Verfahren setzt, erfordert die hinreichende Sicherheit detaillierte Prüfungen von Einzelfällen sowie eine eingehende Analyse der internen Kontrollsysteme. Eine der größten Herausforderungen besteht in der Sicherstellung der Datenqualität und der Überwindung mangelnder Orientierung. Es besteht ein klarer Bedarf an einer weiteren Standardisierung, insbesondere durch die Entwicklung branchenspezifischer Indikatoren zur Bestimmung der Wesentlichkeitsschwellen sowie an einem gezielten Kompetenzaufbau und fachlichen Weiterentwicklung der Prüfer. Diese Schritte sind notwendig, um den steigenden Prüfungsanforderungen gerecht zu werden und die Zuverlässigkeit der offengelegten Informationen für die Berichtsadressaten zu gewährleisten.

Schlagworte: Nachhaltigkeitsberichterstattung, Doppelte Wesentlichkeit, Externe Prüfungsplicht

Keywords: Corporate Sustainability Reporting Directive, double materiality, sustainability assurance

Publikationsstatus: Veröffentlicht in: *Zeitschrift für Umweltpolitik & Umweltrecht (ZfU)* Heft 4 (2024): 583–621

Hinweis: Aufgrund von Formatierungsanpassungen in dieser Dissertation kann es zu geringfügigen Abweichungen gegenüber der veröffentlichten Version kommen.

2.2 Einführung

Im Zuge des Pariser Klimaabkommens wurde eine transparente Berichterstattung über Umweltbelange als Erfolgsfaktor bei der Bekämpfung des Klimawandels deklariert (United Nations Hrsg. 2015, S. 12–14), was angesichts des Verursacherprinzips in puncto negativer Nachhaltigkeitsentwicklung schlüssig scheint (Schaltegger/ Wagner 2011, S. 223). Verpflichtende Offenlegungsvorschriften können das Informationsumfeld von Unternehmen verbessern (Krueger et al. 2021, S. 44). Im Rahmen der Förderung der nachhaltigen Entwicklung von Unternehmen soll durch Einführung der Corporate Sustainability Reporting Directive (CSRD) die Nachhaltigkeitsberichterstattung der Unternehmen auf die Ebene der gegenwärtigen Finanzberichterstattung gestellt werden (RL 2022/2464 ErwG 1, 24). Die CSRD sieht eine gestaffelte Anwendung, beginnend mit dem Geschäftsjahre 2024 vor, wobei im ersten Anwendungsjahr sämtliche Unternehmen betroffen sind, die schon iSd RL 2014/95/EU berichtspflichtig waren. Ab dem Geschäftsjahr 2025 sind nachfolgend sämtliche große Unternehmen bzw. Mutterunternehmen eines Konzerns dazu verpflichtet, einen umfassenden und unabhängig geprüften Nachhaltigkeitsbericht zu veröffentlichen (RL 2022/2464 Art. 1 Abs. 1&13). Allein in Deutschland sind anstelle der bisherigen 500 nun ca. 15.000 Unternehmen direkt (DRSC Hrsg. 2023), und – trotz des Entgegensteuerns der Europäischen Union (EU) – viele weitere Unternehmen indirekt über den sog. „Trickle-Down Effect“ (Schütte/Dickel 2024 S. 864) betroffen.

Die Bedeutungszuschreibung der Wesentlichkeitsanalyse in diesem Kontext ist unumstritten hoch, da sich der Großteil der Berichtsinhalte über jene Analyse bestimmt (ESRS 1.25–1.28). So wird die Wesentlichkeitsanalyse als „Herzstück“ (Baumüller 2024, S. 78) der Nachhaltigkeitsberichterstattung bezeichnet. Unternehmen sind seither verpflichtet zu berichten, wie sich ihr Geschäft auf Menschen und Umwelt auswirkt. Zudem werden Informationen dazu verlangt, welche Auswirkungen die Nachhaltigkeitsaspekte auf das Geschäftsergebnis, ihre Lage und ihren Geschäftsverlauf haben, fortlaufend auch als Prinzip der „doppelten Wesentlichkeit“ bezeichnet (RL 2022/2464, ErwG 29).

Mit der CSRD hat auch eine externe Prüfungspflicht des Nachhaltigkeitsberichts Einzug gefunden (RL 2022/2464 Art. 1 Abs. 13). Vorerst ist eine Prüfung mit begrenzter Sicherheit vorgesehen, wobei mit Etablierung geeigneter Standards der Grad der Prüfungssicherheit an den der finanziellen Berichterstattung angeglichen werden soll (RL 2022/2464 ErwG 60 f.). Der Prüfungsstandard International Standard on Assurance Engagements (ISAE) 3000

(Revised) gilt bis dato als einschlägiger Standard für die Prüfung nichtfinanzieller Informationen (IFAC/AICPA & CIMA Hrsg. 2024, S. 15). Allerdings wurde dieser Standard gem. Rn A8 ISAE 3000 (Revised) nicht explizit für diesen Zweck konzipiert. Im September 2024 wurde der ISSA 5000 verabschiedet, welcher auf dem ISAE 3000 aufbaut und die spezifischen Anforderungen an die Prüfung von Nachhaltigkeitsinformationen adressieren soll (IAASB Hrsg. 2024). Die Wesentlichkeitsanalyse, als zentraler Bestandteil der Nachhaltigkeitsberichterstattung, stellt einen wichtigen Prüfungsgegenstand dar – gleichwohl Ersteller und Prüfer vor neuen Herausforderungen stehen (Freiberg et al. 2023, S. 1110 Rn. 56; Säuberlich/Jordan 2024, S. 140). Prüfer müssen mit Inkonsistenzen und einem niedrigeren Qualitätsniveau des Prüfungsobjekts umgehen können und sich fachlich umfassend weiterbilden (Löffler 2023, S. 342; Disser/Vogl 2023, S. 223).

Die Herausforderungen, mit denen Anwender bei der Umsetzung konfrontiert sind (Meyer/Klose 2023, S. 1244), unterstrich der Bundesjustizminister kürzlich im Rahmen einer Pressekonferenz und merkte überspitzt an, dass die CSRD „quasi eine staatliche Absatzgarantie für Wirtschaftsprüferleistungen“ sei (Buschmann, 2024). Angesichts des Bedarfs an belastbaren Nachhaltigkeitsinformationen (Lanfermann/ Hebestreit 2021, S. 305; Schäfer/Schöneberger 2021, S. 634), der zahlreichen Freiheitsgrade (Velte 2021, S. 620) in der bisherigen Berichtspraxis sowie der daraus resultierenden mangelnden Qualität und Quantität der nichtfinanziellen Berichte (Freiberg/Skoluda 2021, S. 469 f.), zeigt dieser Beitrag expressis verbis, dass die CSRD mehr ist als eine Absatzgarantie für Wirtschaftsprüferleistungen. Vielmehr wird deutlich, dass eine externe Prüfungspflicht notwendig ist. Das Fehlen eines maßgeschneiderten Prüfungsstandards, dringend benötigtes Fachpersonal sowie fehlende Erfahrungswerte im Umgang mit der Prüfung von Nachhaltigkeitsinformationen (Quick 2024b, S. 468; Velte 2023, S. 1290), lässt allerdings die Frage zu, inwiefern die Wirtschaftsprüfung ihrem neuen Aufgabengebiet gerecht werden kann. Im Rahmen einer Befragung von 200 Privatanlegern sah man Wirtschaftsprüfer bspw. als eher ungeeignet für die Prüfung von Nachhaltigkeitsberichten an (Quick 2024, S. 312–315), gleichwohl man aufgrund der allgemeinen Erfahrungswerte in der „Anwendung […] konsistenter Prüfungsmethodologie“ Wirtschaftsprüfern die Aufgabe am ehesten zutraut (Velte 2022, S. 64 f.). Cohen und Simnett (2015) werfen einige Forschungsideen rund um die Prüfung von Nachhaltigkeitsberichten, der Wesentlichkeit sowie der Stakeholder-Einbindung auf. Wir greifen diesbezüglich die Forschungsfrage auf, wie die Prüfung von Nachhaltigkeitsinformationen gestaltet werden kann, um ausreichende Prüfungssicherheit zu erlangen. Wir präzisieren dahingehend den

Untersuchungsgegenstand und analysieren mithilfe der qualitativen Inhaltsanalyse von 12 durchgeführten Experteninterviews das aktuelle Vorgehen von Wirtschaftsprüfern bei der Prüfung der doppelten Wesentlichkeit, um so unter anderem Implikationen für Prüfungsdienstleister, Regulierer und berichterstattende Unternehmen ableiten zu können.

> Forschungsfrage: *Wie kann die Prüfung der doppelten Wesentlichkeitsanalyse gestaltet werden und welche Herausforderungen gilt es dabei zu bewältigen?*

2.3 Entwicklung der doppelten Wesentlichkeitsanalyse

2.3.1 Zunehmende Bedeutung und Regulierung der Nachhaltigkeitsberichterstattung

Spätestens mit der Veröffentlichung des Grünbuchs im Jahr 2001 fällt das Thema Corporate Social Responsibility (CSR) in den Verantwortungsbereich der EU (Europäische Kommission Hrsg. 2001, S. 3). Dieses Anliegen der EU transformierte sich 2011 in eine Strategie (Europäische Kommission Hrsg. 2011) und manifestierte sich mit der Veröffentlichung der Non Financial Reporting Directive (NFRD bzw. RL 2014/95/EU) in obligatorischen Berichtspflichten. Um dem erhöhten Informationsbedarf von Abschlussadressaten und Verbrauchern gerecht zu werden (RL 2014/95/ EU, ErwG 3), eine adäquate bzw. ganzheitliche Risikobeurteilung zu ermöglichen (RL 2014/95/EU, ErwG 5,8), die Vergleichbarkeit nichtfinanzieller Informationen zu verbessern (RL 2014/95/EU, ErwG 6), nachhaltige Investitionsanreize zu stärken (RL 2014/95/EU, ErwG 12), und ein rechtliches Rahmenwerk für die nichtfinanzielle Berichterstattung zu schaffen, welches die Sensibilisierung von Unternehmen in Bezug auf soziale und umweltbezogene Verantwortung adressiert (RL 2014/95/EU, ErwG 2), verabschiedete die EU im Jahren 2014 die Richtlinie 2014/95/EU, welche „[…] die Angabe nichtfinanzieller und die Diversität betreffender Informationen [normiert]“ (RL 2014/95/EU, S. 1). So sind gem. Art. 19 a RL 2013/34/EU und Art. 29 a RL 2013/ 34/EU große Unternehmen sowie Konzerne von öffentlichem Interesse verpflichtet, seit dem Geschäftsjahr 2017 eine nichtfinanzielle Erklärung abzugeben.

Trotz einer hierauf zurückzuführenden qualitativen Verbesserung der ESG-Berichterstattung (Fiechter et al. 2022), sah die Europäische Kommission Anlass für die Überarbeitung der nichtfinanziellen Berichterstattung aufgrund der zunehmenden Wichtigkeit globaler Nachhaltigkeitsbemühungen (RL 2022/2464 ErwG 10 f.) sowie der weiterhin nicht zufriedenstellenden Qualität und eingeschränkten Vergleichbarkeit der bisherigen Offenlegungen (RL 2022/2464 ErwG 13). Die in den Jahren 2017 und 2019 erlassenen Leitlinien für die klimabezogene Berichterstattung verfehlten – ihres unverbindlichen

Charakters wegen – das Ziel, die Nachhaltigkeitsberichterstattung um Qualität und Vergleichbarkeit zu ergänzen (RL 2022/2464 ErwG 37).

Die verdichteten Erkenntnisse aus 20 Studien, welche die nichtfinanzielle Berichterstattung-Praxis deutscher Unternehmen in den ersten Anwendungsjahren untersuchten, weisen auf erhebliche Schwachstellen hin (Velte et al. 2020, S. 1351–1357). Die Ergebnisse stehen im Einklang mit Erkenntnissen aus weiteren Studien zur Qualität der Nachhaltigkeitsberichterstattung in den europäischen Mitgliedsstaaten, wobei neben der divergenten Anwendung von Rahmenwerken und ergo einer mangelnden Vergleichbarkeit von Nachhaltigkeitsinformationen, die Ausgestaltung des Prinzips der doppelten Materialitätsperspektive sowie eine fehlende externe Überprüfung der Berichtsinhalte hervorgehoben werden (Fiandrino et al. 2022, S. 274; DRSC Hrsg. 2021, S. 84; Richter et al. 2021, S. 703–711; Velte et al. 2020, S. 1351–1357). So befähigt der bisherige Auslegungsspielraum in Deutschland – aufgrund der HGB-Auslegung des §289c Abs. 3 – bei der Wesentlichkeitsschwelle auf wichtige Angaben zu nichtfinanziellen Risiken zu verzichten, womit Informationsasymmetrien gefördert werden (Velte et al. 2020, S. 1352). Velte (2020, S. 2038) spricht von einem „[…] Schlupfloch […] der [den] Informationswert der nichtfinanziellen Erklärung de lege lata [aushöhlt] […]“. Zudem war aufgrund der diversen Auslegungspraxis die Vergleichbarkeit von Nachhaltigkeitsinformationen stark eingeschränkt (Baumüller/Sopp 2022, S. 23).

Als Ergebnis der angekündigten Überarbeitungen der NFRD trat am 05. Januar 2023 die CSRD auf EU-Ebene in Kraft (RL 2022/2464). Die entscheidendsten Neuerungen durch die CSRD in dem Vergleich zur NFRD sind: ein erweiterter Anwenderkreis, die externe Prüfungspflicht der offengelegten Informationen, die verpflichtende Orientierung an EU-weiten Berichtsstandards – den European Sustainability Reporting Standards (ESRS) –, die Darstellung in einem elektronischen Berichtsformat und die präzisierte Anwendung einer doppelten Wesentlichkeitsperspektive (Sopp et al. 2022, S. 71). Der erweiterte Anwenderkreis begegnet den Greenwashing-Tendenzen – sprich der Veröffentlichung irreführender Informationen hinsichtlich der ESG-Performance des Unternehmens (Albers/Günther 2010, S. 344) – bei bisher nicht standardisierter freiwilliger Nachhaltigkeitsberichterstattung (Hahn/Lüfls 2013, S. 401). Das zu Beginn dieses Jahres in Kraft getretene Set an 12 sektorübergreifenden ESRS enthält zwei themenübergreifende und zehn themenspezifische Standards. ESRS 1 und ESRS 2 enthalten themenübergreifende, allgemeine Anforderungen und Angaben für die Berichterstattung. Die weiteren zehn ESRS behandeln spezifische Umwelt-, Sozial- und

Governance-Themen, welche in Themen, Unterthemen und Sub-Unterthemen gegliedert sind (ESRS 1.4-8).

2.3.2 Funktionsweise und Umsetzung der doppelten Wesentlichkeitsanalyse gem. CSRD bzw. ESRS

Die doppelte Wesentlichkeitsanalyse ist die Grundlage für die Durchführung der Berichterstattung nach den ESRS, wobei das Ergebnis spezifiziert über welche umweltspezifischen, sozialen und Governance (ESG)-Inhalte ein Unternehmen berichten muss. Dabei ist die Zielsetzung die Berichtsinhalte mithilfe von Relevanzkriterien einzugrenzen. Das Gesamtkonzept beurteilt die jeweiligen Wesentlichkeitsaspekte in eine kurz-, mittel- und langfristige Zeitdimension. Einerseits müssen die Auswirkungen unterschiedlicher Nachhaltigkeitsaspekte auf die finanzielle Situation des Unternehmens selbst berücksichtigt werden („Outside-In“), andererseits über die Wirkung unternehmerischer Tätigkeiten auf das soziale und ökologische Umfeld („Inside-Out“) berichtet werden. Die Berichterstattungspflicht greift demnach schon bei der Relevanz einer dieser Wesentlichkeitsperspektiven und nicht erst bei kumulativer Erfüllung beider. Diese Klarstellung leitet sich aus der heterogenen Auslegung der Wesentlichkeitsperspektive der Mitgliedstaaten ab (ESRS 1.25-28; RL 2022/2464, ErwG 29).

Zwecks Unterstützung der Ersteller von Nachhaltigkeitsberichten wurde ein unverbindlicher Leitfaden für die Umsetzung der Wesentlichkeitsanalyse bereitgestellt. Demzufolge lässt sich der Prozess der Analyse in vier Schritte gliedern, wobei (1) ein Verständnis des Kontextes geschaffen werden muss, (2) die Auswirkungen, Risiken und Chancen (IROs) identifiziert werden müssen, (3) die Bewertung der IROs sowie (4) die Dokumentation des Vorgehens erfolgt (EFRAG Hrsg. 2024, S. 19). Zunächst sollte ein Überblick über das Geschäftsumfeld, Geschäftsaktivitäten, Geschäftsbeziehungen und die relevanten Stakeholdergruppen verschafft werden. Hierzu gehört die Darlegung der strategischen Ausrichtung vor dem Hintergrund des Geschäftsmodells, eine Aufstellung des angebotenen Portfolios von Produkten bzw. Dienstleistungen, der Standorte sowie der vor- und nachgelagerten Wertschöpfungskette (EFRAG Hrsg. 2024, S. 20). Zwecks Identifizierung der IROs sollten die ESRS, und ergänzend die Sammlung zu allen Datenpunkten der ERFAG, welche die Anforderungen aus den jeweiligen ESRS mit den entsprechenden Offenlegungspflichten zusammenfasst, sowie branchenübliche Rahmenwerke herangezogen werden (EFRAG Hrsg. 2024, S. 21). Der dritte Schritt befasst sich mit der exakten Bewertung aller identifizierten IROs und der Bestimmung

ihrer Wesentlichkeit (EFRAG Hrsg. 2024, S. 23). Nachhaltigkeitsthemen der Inside-Out Perspektive werden anhand des Schweregrades, abgeleitet aus dem Ausmaß, dem Umfang und der Unabänderlichkeit, bewertet. Bei Vorliegen einer potenziell negativen Folge für Mensch oder Umwelt muss zusätzlich die Eintrittswahrscheinlichkeit berücksichtigt werden (ESRS 1. 45–46). Die Bewertung der Wesentlichkeit von Nachhaltigkeitsthemen der Outside-In Perspektive ist vergleichbar mit der Wesentlichkeitsbestimmung für die allgemeine Finanzberichterstattung. Zur Bewertung der Chancen und Risiken werden die Eintrittswahrscheinlichkeit und das potenzielle Ausmaß der finanziellen Auswirkung herangezogen (ESRS 1.48-51). Anhand von quantitativen und qualitativen Kriterien sind Wesentlichkeitsgrenzen zu definieren, welche zur Bestimmung der Wesentlichkeit der einzelnen IROs herangezogen werden (ESRS 1.42). Die Dokumentation des Vorgehens wird durch die in den ESRS enthaltenen Angabe-Pflichten formalisiert (ESRS 2.1–3). Grundsätzlich müssen Unternehmen immer über die genutzten Verfahren und Methoden zur Ermittlung der IROs sowie der Bewertung der Wesentlichkeit berichten (EFRAG Hrsg. 2024, S. 25).

2.3.3 Forschungsstand zur Wesentlichkeit im Kontext der Nachhaltigkeitsberichterstattung

Mit Hilfe einer systematischen Literaturanalyse und anschließender Inhaltsanalyse wurden Erkenntnisse aus 90 Studien zur Wesentlichkeitsanalyse im Kontext der Nachhaltigkeitsberichterstattung zusammengetragen. Die in der Literatur verwendeten Forschungsmethoden sind vielfältig, wobei Inhaltsanalysen am häufigsten angewendet werden. Auch die theoretischen Grundlagen der Forschung sind divers: 41 Arbeiten beziehen keine Theorie mit ein, 25 stützen sich auf die Stakeholder-Theorie, 15 auf die Legitimationstheorie und 7 auf die Institutionentheorie. Die meisten Studien konzentrieren sich auf die Bewertung von Nachhaltigkeitsinformationen (34,44 %), während nur wenige Arbeiten wesentliche Themen für Unternehmen und Stakeholder (4,44 %) behandeln. Einfache Wesentlichkeitsdefinitionen, welche einen Investor-bezogenen Fokus legen, führen dazu, dass Unternehmen Angaben zu Risiken, Chancen sowie Auswirkungen bezogen auf die Umwelt vernachlässigen (Fiandrino et al. 2022, S. 672–684). Nutzer von Nachhaltigkeitsberichten haben unterschiedliche Informationsbedürfnisse, was die doppelte Wesentlichkeit adressieren kann. Die Anwendung der doppelten Wesentlichkeit kann die Informationsbedürfnisse der Stakeholder gezielt erfüllen, sofern das Stakeholder-Engagement adäquat berücksichtigt wird (Torelli et al. 2020, S. 480). Ferner steht der Umfang der Berichterstattung oft im Widerspruch zu Präzision und Relevanz (Jørgensen et al. 2022, S. 354 ff.).

Neben der Stakeholder-Theorie erscheint uns die Theorie der Entscheidungsnützlichkeit einschlägig, wonach Informationen dann relevant sind, wenn sie für die Entscheidungen der Adressaten als hilfreich angesehen werden (Staubus 2003, S. 164). Das International Sustainability Standards Board (ISSB) hat zum Ziel, die Berichterstattung dem Informationsbedarf der Berichtsadressaten anzupassen und somit „nützliche" Informationen bereitzustellen (IFRS Foundation Hrsg. 2023, S. 23). Auch die Berichtspflichten nach CSRD sind demgemäß ausgerichtet (RL 2022/2464 ErwG 9, 14, 15, 54). Ob eine Information als Nutzengewinn wahrgenommen wird, hängt dabei vom Informationsbedarf des Adressaten ab (Sterzel 2011, S. 65). Ein höheres Maß an Transparenz allein reicht nicht aus, sondern sollte mit einer unabhängigen Prüfung einhergehen, um die Glaubwürdigkeit sowie die Relevanz und Nützlichkeit der Nachhaltigkeitsberichterstattung für die Stakeholder sicherzustellen (Defond/Zhang 2014, S. 199).

2.4 Prüfung der doppelten Wesentlichkeitsanalyse im Kontext wissenschaftlicher Theorien

2.4.1 Externe Prüfungspflicht von Nachhaltigkeitsberichten und damit einhergehende Herausforderungen

Bisher oblag die Pflicht zur inhaltlichen Prüfung der nichtfinanziellen Erklärung gem. §171 Abs. 1 Satz 4 AktG dem Aufsichtsrat, welcher diese Verantwortung zunehmend im Rahmen der Beauftragung einer externen freiwilligen Prüfung weitergab. Die externe Prüfung wurde in Deutschland fast ausschließlich durch Wirtschaftsprüfer unter Anwendung des Prüfungsstandards ISAE 3000 (revised) durchgeführt (Velte et al. 2020, S. 1357; Philipps 2024).

Um dem Ziel nachzukommen, die Nachhaltigkeitsberichterstattung auf die Ebene der Finanzberichterstattung zu stellen, müssen sich Unternehmen nun einer externen Prüfung des CSRD-Berichts unterziehen. Dabei soll der Bericht zunächst mit begrenzter Sicherheit („limited assurance") und künftig mit hinreichender Sicherheit („reasonable assurance") geprüft werden, was zur Zielerreichung der Niveaugleichheit von Finanz- und Nachhaltigkeitsberichterstattung beitragen soll (RL 2022/2464 ErwG 60 f.). Artikel 34 RL 2013/34/EU sieht hierfür ein Mitgliedstaatenwahlrecht vor, welches die materielle Prüfung des Nachhaltigkeitsberichts durch den Abschlussprüfer, einen weiteren Wirtschaftsprüfer oder andere unabhängige Prüfer außerhalb des Berufsstandes erlaubt.

Der bisherige Regierungsentwurf sieht vorerst von der Nutzung des Wahlrechts ab und definiert gem. § 324e HGB-E die Prüfung der Nachhaltigkeitsberichterstattung als Vorbehaltsaufgabe für Wirtschaftsprüfer, wobei sich die Regierung eine erneute Prüfung dieser Entscheidung vorbehält (Bundesregierung Hrsg. 2024, S. 144–145, 176). Da das Wahlrecht zur Beauftragung alternativer Prüfungsdienstleister in der Kritik steht (Quick et al. 2022, S. 37; Velte 2021, S. 618), welche von empirischen Erkenntnissen begleitet wird, die auf stärkere positive Wirkungen von Prüfung hindeuten, wenn diese durch Wirtschaftsprüfer durchgeführt werden (Quick et al. 2022, S. 41 f.), wird die bisherige Auslegung des Gesetzesentwurfs begrüßt (Quick 2024b, S. 468). Um die hohe Qualität der Prüfungsleistung einhalten zu können, müssten die externen Dienstleister denselben strengen Regeln unterworfen werden wie dem Berufsverband. Diese regulatorischen Rahmenbedingungen sind – zumindest in Deutschland – noch nicht gegeben (Bundesministerium der Justiz Hrsg. 2024, S. 2).

Eine Einbindung in die nationale Enforcement-Prüfung ist zu begrüßen, da so die Glaubwürdigkeit und suggerierte Qualität der publizierten Nachhaltigkeitsberichte erhöht wird, gleichwohl die ausgeweiteten Kontrollpflichten zuständige Institutionen vor große Herausforderungen stellen. Die vorläufige Begrenzung der Prüfungssicherheit schränkt zwar die Belastbarkeit der Nachhaltigkeitsberichterstattung vorerst ein, erscheint aber vor dem Hintergrund des Verhältnismäßigkeitsprinzips plausibel. Strategisch bietet sich demnach die Steigerung der Prüfungsintensität an, um die Reputation von ESG-Informationen zu erhöhen, um so „einen Gleichschritt zur Finanzberichterstattung zu erzielen“ (Velte 2021, S. 617–620). Für diese Gleichstellung bedarf es ausgereifter und standardisierter Prozesse bei den berichterstattenden Unternehmen und Prüfern, weshalb die sukzessive Erhöhung der verlangten Prüfungssicherheit als geeignete und realisierbare Übergangslösung erscheint (WPK Hrsg. 2021, S. 3; Lanfermann/Glöckner 2020, S. 1228 f.). Ferner wird aufgrund des mangelnden Angebots qualifizierter Prüfer sowie an Ermangelung geeigneter Standards bezweifelt, ob die Prüfung mit hinreichender Sicherheit überhaupt schon realisiert werden könnte (Quick 2024b, S. 468). So geht die Prüfung von ESG-Daten mit Herausforderungen einher, da diese oft subjektiv und auf internen Informationen basieren, die nur schwer mit anderen Unternehmen vergleichbar sind. Ferner werden ESG-Daten häufig in Systemen erfasst, die außerhalb der doppelten Buchführung und manchmal auch außerhalb der Unternehmensgrenzen liegen, wie beispielsweise Daten zur Lieferkette. Ferner erfordert die Überprüfung von ESG-Angaben oft spezifisches Fachwissen aus technischen oder wissenschaftlichen Bereichen, das über klassische Rechnungslegungskenntnisse hinausgeht (Krueger et al. 2021, S. 37). Obwohl

Wirtschaftsprüfer nicht immer über diese spezialisierte Expertise verfügen, kann ihre Prüfungserfahrung vieles kompensieren. Insbesondere Prozessverständnis und Erfahrung mit der Identifizierung sowie Bewertung von Kontrollen sind dabei von Bedeutung (Knechel 2021). Die EU-Kommission sieht vor, auf die Nachhaltigkeitsberichterstattung abgestimmte Prüfungsstandards zu verabschieden. Diese Ankündigung ist zu begrüßen, da die Bestätigungsvermerke von Wirtschaftsprüfern nicht per se als aussagekräftiges Qualitätssiegel wahrgenommen werden dürfen, sondern angewandte Prüfungspraktiken maßgeblich für die Transparenz und Verlässlichkeit der Informationen sind (Garcia-Sanchez et al. 2022, S. 103; Boiral/Heras-Saizarbitoria 2020, S. 12).

Entwicklung einschlägiger Prüfungsstandards

Der ISAE 3000 (Revised) wird momentan als Standard zur Prüfung der Nachhaltigkeitsberichterstattung herangezogen (IFAC/AICPA & CIMA Hrsg. 2024, S. 15). Unter eine Prüfung durch den ISAE 3000 (Revised) fallen sonstige betriebswirtschaftliche Prüfungen, welche nicht Teil einer Prüfung von vergangenheitsorientierten Finanzinformationen sind (ISAE 3000 (Revised) Rn. 1). Demgemäß ist der Standard zwar nicht explizit für die Prüfung von Nachhaltigkeitsberichten verfasst worden, kann aber für solche verwendet werden (ISAE 3000 (Revised) Rn. A8). Mit Blick auf die Nutzung des ISAE 3000 (Revised) zur Prüfung von Nachhaltigkeitsberichten hat das IAASB 2021 einen Leitfaden veröffentlicht, welcher explizite Hinweise für die Prüfung von Nachhaltigkeitsinformationen enthält (IAASB Hrsg. 2021, S. 4 Rn. 1–3). Der International Standard on Sustainability Assurance (ISSA) 5000 soll als internationaler Prüfungsstandard ab Ende 2024 für Prüfungen mit begrenzter und hinreichender Sicherheit anzuwenden sein. Er ist die Reaktion auf die Forderung eines Standards einzig für die Prüfung von Nachhaltigkeitsberichten (IAASB 2024). Der Standard kann sowohl von Wirtschaftsprüfern als auch von unabhängigen Prüfungsdienstleitern angewendet werden (IAASB Hrsg. 2023, S. 8 Rn. 14). Da der ISAE 3000 (Revised) als Ausgangsbasis für den ISSA 5000 genutzt wurde, liegen die Inhalte nah beieinander (ED ISSA 5000 Rn. 17). Im direkten Vergleich zum ISAE 3000 (Revised) stellt der ISSA 5000 höhere Anforderungen an die Risikobeurteilung und Wesentlichkeitsbeurteilung (Disser/Skoluda 2024, S. 73).

2.4.2 Theoretische Begründung der externen Prüfungspflicht

Eine systematische Literaturanalyse von 142 Studien stellt 32 verwendete Theorien im Kontext der Offenlegung von Nachhaltigkeitsinformationen heraus. Die einschlägigsten Ansätze –

gemessen an den 22 meist zitiertesten Studien – ist die Stakeholder-Theorie, gefolgt von der Legitimitäts-, Institutions-, Agency- und Signaling-Theorie, wobei insbesondere die Stakeholder- und Legitimitätstheorie kombiniert werden (Del Gesso/Lodhi 2025). Vor dem Hintergrund der Prüfungspflicht stützen wir uns ähnlich wie Quick (2024b, S. 465 ff.) auf die Prinzipal-Agent-Theorie und er-läutern in diesem Kontext zwei Beziehungen, die von Informationsasymmetrien geprägt sein können. Wir ergänzen diesen Ansatz mit der weiter oben eingeführten Theorie der Entscheidungsnützlichkeit. Wir postulieren, dass Informationsasymmetrien zwischen dem berichterstattenden Unternehmen und den Stakeholdern bestehen können.

Die Prinzipal-Agent-Theorie befasst sich mit Konflikten zwischen Agenten und Prinzipalen, die auf Informationsasymmetrien zurückzuführen sind. Aufgrund der nutzenmaximierenden Eigenschaften der Vertragsparteien und der Tatsache, dass der Agent einen Informationsvorsprung hat, besteht die Gefahr, dass der Agent nicht immer im besten Interesse des Prinzipals handelt (Manne 2005 S. 170; Manne 1996, S. 155 f.; Jensen/Meckling 1976, S. 308). Gemäß ErwG 9 benennt die EU Fremdkapitalgeber und Akteure der Zivilgesellschaft als zwei Hauptnutzergruppen von Nachhaltigkeitsinformationen. Greift man die in Erwägung stehenden Gründe für die Verschärfung der Berichtspflichten auf, lässt sich das Problem der Informationsasymmetrie zwischen Berichtsadressaten und berichterstattenden Unternehmen feststellen (RL 2022/2464 ErW 13–15). So sind die Stakeholder des Unternehmens aufgrund unzureichender Transparenz und mangelnden Einblicks in die internen Prozesse nicht in der Lage, die Handlungen des Unternehmens vollständig nachzuvollziehen („Hidden Action“). Dies birgt das Risiko, dass die Unternehmensleitung ihre überlegenen Informationen zum eigenen Vorteil verwendet. Dieser potenzielle Interessenkonflikt wird als „moral hazard“ bezeichnet und beschreibt das Problem, dass die Unternehmensleitung versucht, ihren persönlichen Nutzen zu maximieren, während die Stakeholder dies aufgrund fehlender Informationen nicht kontrollieren können. Überträgt man dies auf expliziten Anwendungsfall, so könnten Unternehmen bspw. wesentliche finanzielle Risiken oder negative Auswirkungen auf die Umwelt vorenthalten. So beschrieben Yu et al. (2020), dass Stakeholder Schwierigkeiten haben glaubwürdige Nachhaltigkeitsinformationen zu erhalten und somit Investitionshürden aufgebaut werden. Nachhaltigkeitsberichte – insbesondere verpflichtende (Chung et al. 2024) – können dieses Problem adressieren und Informationsasymmetrien abbauen (Reineke 2023, S. 688; Cui et al. 2018, S. 568; Zorio et al. 2013, S. 497 f.), wobei weitere Erkenntnisse zeigen, dass die Veröffentlichung allein nicht ausreicht. Die Belastbarkeit

bzw. Glaubwürdigkeit der Informationen spielt insbesondere für Kapitalgeber eine wichtige Rolle (Clarkson et al. 2019, S. 2468 f.; Reimsbach et al. 2018, S. 559 ff.), welche ungeprüften Nachhaltigkeitsinformationen keinen signifikanten Wert beimessen (Clarkson et al. 2019, S. 2468 f.). Insbesondere für das *moral hazard* Problem scheint die Prüfung der Nachhaltigkeitsberichterstattung als geeignetes Instrument. Die Prüfung ist als Monitoring zu verstehen, um das tatsächliche Verhalten des Agenten – hier das Unternehmen bzw. die Unternehmensleitung – zu überwachen und die Glaubwürdigkeit der Informationen zu erhöhen (RL 2022/2464 ErW 60). Die Prüfung von ESG-Berichten erhöht die Qualität der Berichterstattung, wobei dieser Effekt zunimmt, sofern Wirtschaftsprüfer die Prüfung vornehmen (Ballou et al. 2018, S. 185; Zorio et al. 2013, S. 498). Die Prüfung mit hinreichender Sicherheit durch einen Wirtschaftsprüfer kann also als wirksames Instrument für die Genauigkeit und die Glaubwürdigkeit von Nachhaltigkeitsinformationen dienen und somit Informationsasymmetrien abbauen. In Ländern mit einer höheren Stakeholder Orientierung ist dieser Effekt auch unabhängig von der Wahl des Prüfers und der Prüfungssicherheit zu beobachten (Cuadrado-Ballesteroz et al. 2017, S. 1149, 1154.) Die Wesentlichkeitsanalyse und deren Prüfung spielt demgemäß eine zentrale Rolle (Nishitani et al 2020, S. 12), da somit sichergestellt wird, dass sämtliche für die Stakeholder relevanten Informationen offengelegt werden (ESRS 1.41). Unter Bezugnahme der Relevanzkriterien wird somit die Vollständigkeit des Nachhaltigkeitsberichts adressiert, wobei eine externe Prüfung zur tatsächlichen Vollständigkeit von Nachhaltigkeitsberichten beitragen kann (Weuster et al. 2020, S. 1295).

2.5 Empirisches Forschungsdesign

2.5.1 Methode der qualitativen Experteninterviews

Aufgrund des sukzessiven Umsetzungsauftrags iSd RL 2022/2464 Art. 5 Abs. 2 sind die ersten Nachhaltigkeitsberichte gem. CSRD sowie deren korrespondierenden Prüfungsvermerken im ersten Quartal 2025 zu erwarten. Die geprüften Berichte der großen – nicht kapitalmarktorientierten – Unternehmen liegen hingegen erst ab 2026 zur Analyse vor, was zu dem jetzigen Zeitpunkt eine quantitative Forschungsmethode ausschließt. Für die Erhebung von Informationen zum Umsetzungsstand der Wesentlichkeitsanalyse sowie dem Vorgehen bei der Prüfung dieser, bietet sich die Nutzung einer qualitativen Forschungsmethode – wie der Einsatz von Experteninterviews – an, da so einerseits theoriegeleitete Problemstellungen (Mayring 2016, S. 71), als auch praktische Problemstellungen adressiert und zielgerichtete Implikation für die Praxis bereitgestellt werden können (Gläser-Zirkuda et al. 2022, S. 238).

Leitfadengestützte Experteninterviews zeichnen sich durch einen strukturierten Gesprächsverlauf aus und sind eine häufig verwendete Form der qualitativen Datenerhebung (Misoch 2019; S. 125; Helfferich 2022, S. 875). Diese Form unterscheidet sich von anderen Interviewformen durch die eindeutige Bezugsgruppe der Interviewpartner – den Experten. Als Experten werden Personen angesehen, die eine besondere Kompetenz bzw. spezielles Wissen innerhalb eines Themengebiets aufweisen. Diese Expertise geht regelmäßig mit der Berufsrolle einher (Niederberger 2015, S. 40; Misoch 2019, S. 127). Mayring differenziert zwischen stark und schwach strukturierten Interviewtechniken. Für die vorliegende Forschungsarbeit wurde die Methode des semistrukturierten Interviews – auf Grundlage eines zuvor entwickelten Interviewleitfadens – eingesetzt. Diese Methode bietet einerseits die Möglichkeit einer strukturierten und vergleichbaren Auswertung (Kühl et al. 2009, S. 35; Misoch 2019, S. 66) und andererseits die Flexibilität auf die Antworten der Interviewpartner unmittelbar einzugehen. Dadurch können unklare Aussagen geklärt oder wichtige Aspekte durch gezieltes Nachfragen vertieft werden (Mayring 2016, S. 67). Der Interviewleitfaden wurde auf Grundlage der in den vorangegangenen Kapiteln gewonnenen Erkenntnisse entwickelt und in Übereinstimmung mit der Forschungsfrage strukturiert. Zwecks Validierung der Verständlichkeit und Vollständigkeit sowie dem Ausloten potenzieller Verbesserungen, wurde der entwickelte Leitfaden mit einem Experten aus der Wirtschaftsprüfung einem Pre-Test unterzogen (Mayring 2016, S. 69; Weichbold 2022, S. 443).

2.5.2 Auswahl der Interviewpartner*innen und Interviewdurchführung

Die Arbeit soll einen Einblick in die Prüfung der Wesentlichkeitsanalyse gewährleisten, davon ausgehend einen Orientierungsrahmen für die Praxis bieten, sowie etwaige Implikationen für Regulierer ableiten. Status Quo ist im Referentenentwurf des Umsetzungsgesetzes die Prüfung der Nachhaltigkeitsberichterstattung als Vorbehaltsaufgabe von Wirtschaftsprüfern vorgesehen (Bundesregierung Hrsg. 2024, S. 144–145, 176). Die Datenerhebung fand demgemäß mit zwölf Experten aus der Wirtschaftsprüfung statt, die anhand ihrer Stellenbezeichnung als Experten für die Nachhaltigkeitsberichterstattung identifiziert wurden. Zwei von ihnen arbeiteten bei einer der sog. Big4-Gesellschaften, acht Personen bei den sog. Next10-Gesell-schaften und zwei der Interviewpartner bei kleineren Kanzleien. Insgesamt haben 12 von 30 angefragten Personen einem Interview zugestimmt.

Nr.	Position	Berufstitel	Big 4 /Next Ten/Klein
I.1	Director	WP (SA)	Klein
I.2	Manager Sustainability Assurance	Consultant	Big4
I.3	Partner	WP	Next Ten
I.4	Senior Manager	WP	Next Ten
I.5	Partner	WP	Next Ten
I.6	Director	WP (SA)	Klein
I.7	Director	WP (SA)	Next Ten
I.8	Senior Manager	WP	Next Ten
I.9	ESG Consultant	Consultant	Big 4
I.10	Senior Manager	WP	Next Ten
I.11	Partner	WP	Next Ten
I.12	Partner	WP	Next Ten

Tabelle 2.1: Angaben zu den Interviewpartnern

2.5.3 Datenauswertung

Die Auswertung der Interviews wurde anhand einer qualitativen Inhaltsanalyse vorgenommen. Dies ist eine methodisch empirische Auswertungsmethode, um transkribierte Ausgangsdaten systematisch zu analysieren (Weber/Wernitz 2021, S. 6 f.) mit dem Ziel aus der erhobenen Datengrundlage eine logische Struktur herauszuarbeiten, die auf die Fragestellung der Forschung ausgerichtet ist (Gläser-Zikuda et al. 2022, S. 241–244). Demnach wird die qualitative Inhaltsanalyse als ein systematisches, regelgeleitetes Verfahren zur Auswertung von textbasierten Daten beschrieben, das sowohl deduktive und induktive Kategorienbildung als auch qualitative und quantitative Elemente umfassen kann (Mayring/Fenzl 2022, S. 691–704). Der gesamte dargestellte Analyseprozess wurde durch die Software MAXQDA unterstützt (Kuckartz/Rädiker 2024, S. 57–75). Die Auswertung der Daten folgt den skizzierten sechs Schritten.

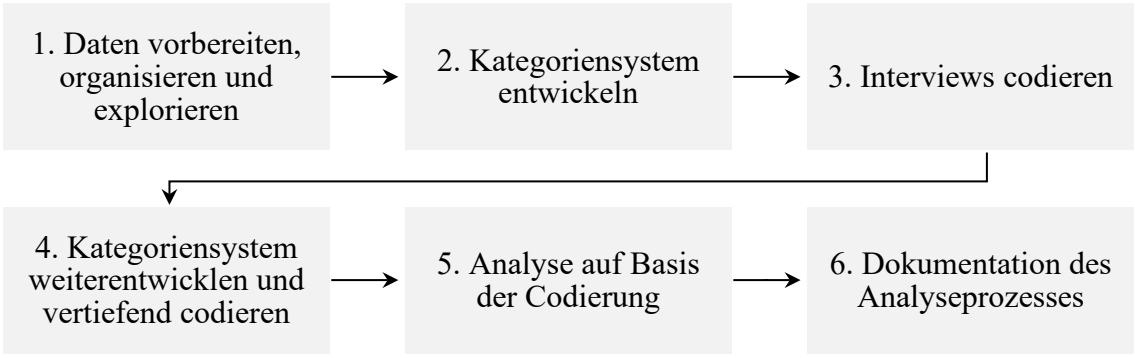


Abbildung 2.1: Datenauswertung (In Anlehnung an Kuckartz/Rädiker 2024)

Die Exploration der Interviews erfolgt im Anschluss an die Transkription. Hierzu wurden die Interviews aufmerksam gelesen und erste Auffälligkeiten und Muster hervorgehoben. Mittels Exploration konnte ein erster Eindruck gewonnen werden, was einzelne Interviewpartner nicht geäußert haben und wo erwartete sowie unerwartete Zusammenhänge bestehen. Diese erste inhaltliche Auseinandersetzung mit den einzelnen Interviews gewährleistet einen ganzheitlichen Überblick über das Datenmaterial und wurde mittels Kurz-Zusammenfassungen dokumentiert (Kuckartz/ Rädiker 2024, S. 14–19). Anschließend wurde eine Paraphrasierung der einzelnen Interviewaussagen vorgenommen. Dazu wurden die Textstellen in eigenen Worten wiedergegeben und kürzer gefasst, um die Bedeutung der Aussagen komprimiert darzustellen (Kuckartz/Rädiker 2024, S. 20 f.). Das Kategoriensystem wurde mittels Kombination einer induktiven und deduktiven Kategorienbildung entwickelt. Vor Beginn der Analyse wurden anhand der theoretischen Grundlage erste Kategorien abgeleitet und für die Strukturierung des Interviewleitfadens genutzt (Mayring/Fenzl 2022, S. 699). Im Rahmen des „Basiscodieren“ wurde eine erste Kategorienzuweisung der einzelnen Interviewaussagen vorgenommen. Hinsichtlich einzelner Aussagen, die keiner deduktiv entwickelten Kategorie eindeutig zugeordnet werden konnten, wurde das Kategoriensystem weiterentwickelt (Kuckartz/Rädiker 2024, S. 45–54). Die so codierten Interviewaussagen werden in dem Rahmen der „Feincodierung“ je Kategorie betrachtet. Die Aussagen je Kategorie wurden inhaltlich als Ganzes analysiert und durch Subkategorien weiter differenziert (S. 62–63). Den Abschluss der qualitativen Inhaltsanalyse bildet die Ergebnisdarstellung anhand der vorgenommenen Analyse und Kategorienbildung. Das entwickelte Kategoriensystem ist der nachfolgenden *Abbildung* zu entnehmen, wobei die grau hinterlegten Kategorien induktive gebildet wurden.

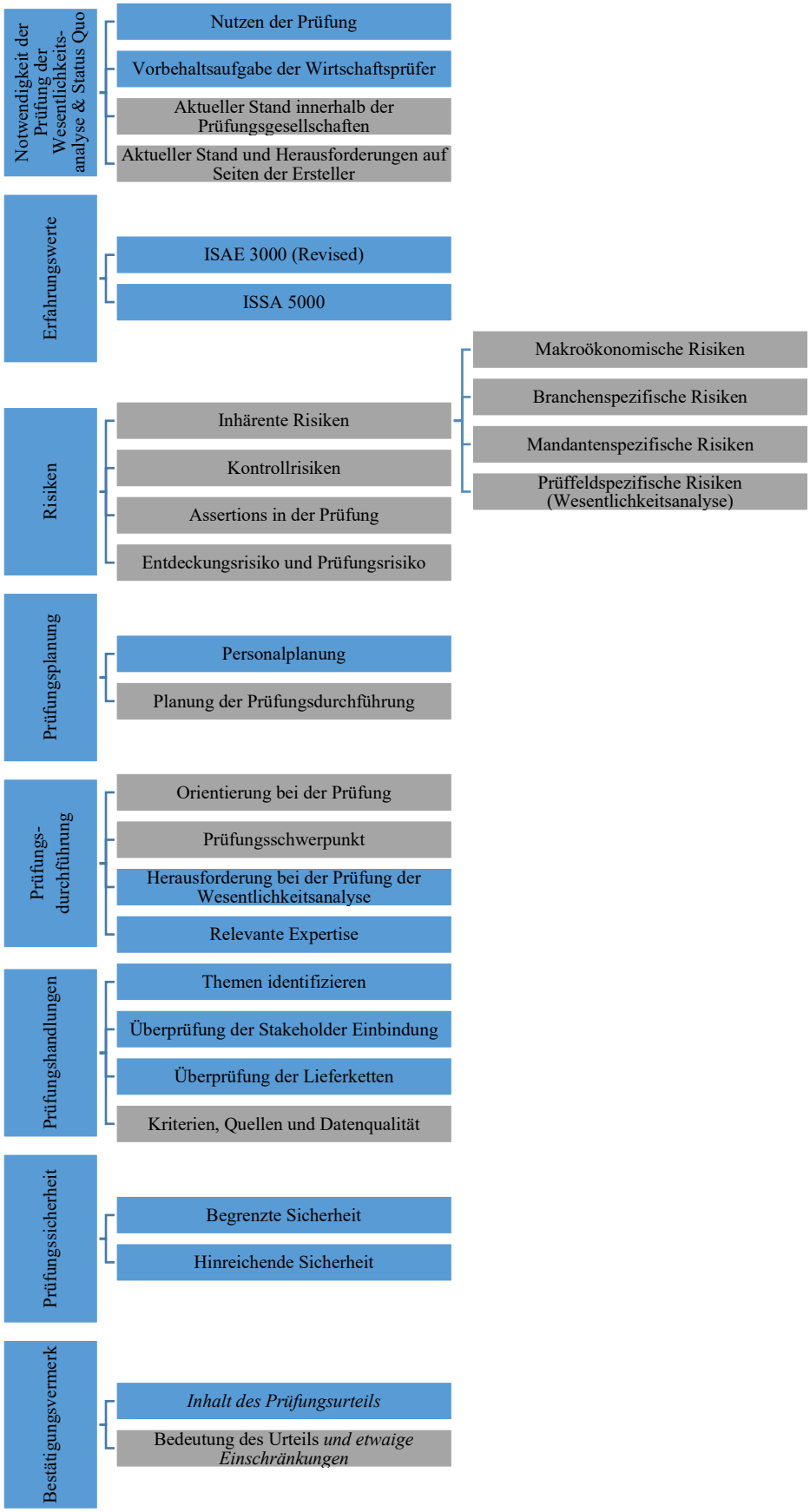


Abbildung 2.2:Ausschnitt des Kategoriensystems

2.6 Ergebnisse der empirischen Untersuchung

2.6.1 Notwendigkeit der Prüfung & Status quo

Zunächst sahen einige der Interviewpartner die Notwendigkeit der Prüfung als normativ gegeben an (Interviewpartner 3, 4, 7, 11). Ungefähr die Hälfte der Befragten sprachen der Prüfung der Wesentlichkeitsanalyse eine hohe Relevanz zu (Interview-partner 1–2, 4, 8–9). Interviewpartner 5 und 6 zur Folge könnten ohne eine Überprüfung, einerseits wesentliche Themen unterschlagen oder nur selektiv veröffentlicht werden, anderseits zu viele unwesentliche Themen den Bericht verwässern. So komme es bei freiwilligen Berichten bis dato noch oft dazu, den Bericht mit unwichtigen Dingen zu überladen (Interviewpartner 11). Die Verlässlichkeit der Informationen als auch die Güte und Schärfe der Analyse seien daher ausschließlich durch eine externe Prüfung zu gewährleisten (Interviewpartner 5, 6, 11).

Nutzen der Prüfung

Das Urteil zur Nachhaltigkeitsberichterstattung kann Interviewpartner 9 zur Folge ein vertrauenswürdiges Signal an Investoren darstellen und dazu beitragen, dass die Berichterstattung an Reputation gewinnt sowie das Vertrauen in Institutionen. Ein weiterer Interviewte hob insbesondere die durch die Prüfung erlangte Verlässlichkeit für Stakeholder hervor (Interviewpartner 4). Interviewpartner 3 hob neben der legitimierenden Wirkung des Testats, welches dem Markt signalisiert gesetzliche Anforderungen zu erfüllen, die gesellschaftsspezifischen Auswirkungen hervor. So könne das Testat Einfluss auf Gesellschafterstruktur und die Finanzierung (so auch Interviewpartner 9) von Unternehmen haben. Ferner sei die Prüfung ein wichtiger Hebel, um die Transformationsziele der Politik zu unterstützen (Interviewpartner 6, 7, 9), gleichwohl zahle diese aber auch auf das Eigeninteresse der Unternehmen ein, da durch eine kontinuierliche Validierung und Überprüfung der wesentlichen Risiken und Chancen strategische Überlegungen abgeleitet werden können. Die geschaffene Transparenz, durch eine Prüfung, sollte von den Unternehmen daher auch genutzt werden (Interviewpartner 9). Außerdem würden die berichtenden Unternehmen durch die Prüfung herausgefordert werden, ihre eigenen Ziele und Maßnahmen ausgehend von der Wesentlichkeitsanalyse zu hinterfragen und anzupassen (Interview-partner 1).

Interviewpartner 12 hingegen kritisiert die fehlende Verhältnismäßigkeit und den pauschalen Anwendungskreis, wonach ein Großteil der Berichterstattungspflicht auf „kleinere“ Unternehmen fällt, obwohl deren Beitrag zum Carbon Footprint minimal sei. Bei dieser

Argumentation berief er sich auf eine Studie der Universität Köln. Ferner werde der Nutzengewinn der Stakeholder somit ineffizient adressiert.

Vorbehaltsaufgabe der Wirtschaftsprüfer

Insgesamt sprachen sich fünf der Interviewten für die Vorbehaltsaufgabe der Prüfung aus (Interviewpartner 4–6, 7, 9), wohingegen sich der Rest hierzu nicht positionierte. Einer der Interviewten führte die Zustimmung zur Vorbehaltsaufgabe auf ein höheres Prüfungsniveau durch den Berufsstand zurück (Interviewpartner 5). Zwei der Experten begründeten ihren Zuspruch, indem Sie auf die Erfahrungswerte mit Prüfungen verwiesen und die allgemein geltenden Qualitätssicherungsmaßnahmen und Haftungsbedingungen für Wirtschaftsprüfer anführten, die es in anderen Prüfungsinstitution so nicht geben würde (Interviewpartner 5, 7). Gleichwohl sehen Sie, dass spezifische technische Expertise, wie im Bereich von Umweltgutachten oder anderen Teilaspekten, weiterhin von Institutionen wie dem TÜV oder Umweltingenieuren beigesteuert werden kann, ähnlich wie bei der Hinzuziehung von Versicherungsmathematikern bei Pensionsrückstellungen (Interviewpartner 3, 5–7). Interviewpartner 4 hob die Synergieeffekte – wie die Kenntnis potenziell relevanter Kontrollen – hervor, die entstehen würden, wenn der Abschlussprüfer auch den Nachhaltigkeitsbericht prüft.

Aktueller Stand innerhalb der Wirtschaftsprüfungskanzleien

In den vertretenen Gesellschaften ist es üblich, dass eine Expertengruppe existiert, welche die Prüfung der Nachhaltigkeitsberichterstattung betreut bzw. begleitet und in der Vergangenheit bereits freiwillige Prüfungen durchgeführt hatte (Interviewpartner 1, 2, 6, 7, 10, 12). Nahezu alle Prüfer (Interviewpartner 1–9, 12) sind Teil eines solchen Experten-Teams zur Nachhaltigkeitsberichterstattung oder hatten zumindest im Rahmen ihrer Prüfertätigkeit Kontakt mit Nachhaltigkeitsthemen (Interviewpartner 10, 11). 3 von 10 der Wirtschaftsprüfer gaben an, bereits an Fortbildungen teilgenommen zu haben und die Zusatzqualifikation Sustainability-Auditor des IDW (IDW Hrsg. 2024) abgeschlossen zu haben (Interviewpartner 1, 6, 7). Momentan wird laut Interviewpartner 4, 9 und 10 ein Personalzuwachs in den Kanzleien forciert, um die zusätzlichen Prüfungsaufgaben stemmen zu können. Allerdings erschwere der vorherrschende Fachkräftemangel dieses Vorhaben (Interviewpartner 9, 10).

Einige der Interviewpartner hatten zu dem Zeitpunkt kaum praktische Erfahrung in der Prüfung der Wesentlichkeitsanalyse im Sinne der ESRS gesammelt (Interview-partner 3, 4, 6, 8). Die theoretische Auseinandersetzung mit der Thematik hat jedoch

bei den meisten Befragten bereits begonnen (Interviewpartner 1, 2, 5, 7, 11, 12). Einige unterstützen Mandanten bei der Erstellung ihrer Wesentlichkeitsanalysen im Rahmen von Beratungsprojekten (Interviewpartner 2, 3, 12). Ferner sind bereits erste Review-Tätigkeiten durchgeführt worden (Interviewpartner 11, 12).

Aktueller Stand und Herausforderungen auf Seiten der Ersteller

Die Erfahrungen, welche die Gesprächspartner bereits mit ihren Mandanten gesammelt hatten, zeigten ein sehr heterogenes Bild (Interviewpartner 1, 3, 7–9). Einige Mandanten befassten sich bereits länger mit der Erstellung der Wesentlichkeitsanalyse, andere hatten noch nicht mit einer Auseinandersetzung begonnen (Interviewpartner 1, 2). Bereits erarbeitete Wesentlichkeitsanalysen sind laut Interviewpartner 2 und 3 stellenweise fehlerhaft und es bestehen Verständnisprobleme bezüglich der Erstellung. Zeit- und Kostenfaktoren werden von Interviewpartner 6 als Grund hierfür angeführt. So sind berichtspflichtige Unternehmen ebenfalls mit dem Problem des Personalmangels konfrontiert und haben ergo Schwierigkeiten, Kapazitäten einzuräumen (Interviewpartner 8). Ein kleinerer Teil der Mandanten zeigt zudem Unverständnis für die gesetzlichen Anforderungen (Interviewpartner 6). Erstanwender, die für das Geschäftsjahr 2024 berichten müssen, waren zum Zeitpunkt der Interviewdurchführung hingegen bei der Finalisierung des Nachhaltigkeitsberichts (Interviewpartner 11). Insbesondere kleinere Unternehmen verfügen laut Interviewpartnern 1, 4, 5 und 10 oft über begrenzte finanzielle Mittel und hätten Schwierigkeiten, kurzfristig qualifiziertes Personal für die Nachhaltigkeitsberichterstattung zu rekrutieren.

2.6.2 Erfahrungswerte der Prüfer und einschlägige Standards

Die Teilnehmenden wurden zu ihren bisherigen Erfahrungen mit gängigen Prüfungsstandards für die Prüfung der Nachhaltigkeitsberichterstattung gefragt. Dabei standen vor allem der Prüfungsstandard ISAE 3000 (Revised) und der Entwurf des ISSA 5000 im Fokus.

Der ISAE 3000 (Revised) wird für Prüfungen von Nachhaltigkeitsberichten genutzt, aber auch für prüferische Durchsichten unabhängig von Nachhaltigkeitsthemen (Interviewpartner 4, 7, 10, 12). Vier der Interviewpartner deuteten an, dass sie keine praktische Erfahrung mit dem ISAE 3000 (Revised) haben (Interviewpartner 1, 3, 5, 8). Interviewpartner 9 verdeutlichte, dass eine umfassende inhaltliche Auseinandersetzung mit Prüfungsstandards insbesondere durch die Grundsatzabteilung der Gesellschaft stattfindet und man sich bei der Durchführung der Prüfung auf Praxishinweise dieser stützt. Die Mehrheit der Experten, denen der ISAE 3000 (Revised) aus der Prüfungspraxis gut bekannt ist, weisen auf einen Verbesserungs- und

Präzisierungsbedarf hin (Interviewpartner 2, 4, 9, 10, 11). So wird der Standard als „Ausweich-Standard“ (Interviewpartner 4) oder „Notlösung“ (Interviewpartner 10) bezeichnet, welcher nicht die spezifischen Anforderungen der Prüfung von Nachhaltigkeitsberichten adressiere (Interviewpartner 7).

Der Entwurf des ISSA 5000 war allen Interviewpartnern bekannt, auch wenn noch keine praktischen Erfahrungswerte vorliegen. Es wurde deutlich, dass die Vorgaben der unternehmensinternen Fachabteilungen relevanter seien, solange der ISSA 5000 nicht final umgesetzt sei (Interviewpartner 9). Einige der Gesprächsteilnehmer hatten bereits mit ersten theoretischen Auseinandersetzungen und Diskussionen zur Umsetzung des ISSA 5000 begonnen (Interviewpartner 7, 12). Im Zuge der Veröffentlichung des Entwurfs zu dem ISSA 5000 hat Interviewpartner 7 einen Abgleich zwischen dem ISAE 3000 (Revised) und dem ISSA 5000 vorgenommen. Die Erkenntnisse des Abgleichs wurden unter anderem genutzt, um Konkretisierungen des ISAE 3000 (Revised) für laufende Prüfungen vorzunehmen. Durch den Standard erhoffen sich die Interviewpartner konkrete Vorgaben für die Prüfung (Interviewpartner. 7, 10). Einer der Interviewpartner erwartet, dass durch das Inkrafttreten des ISSA 5000, sowohl für die Wirtschaftsprüfung als auch für den Mandanten, mehr Orientierung geschaffen werde und Zweifelsfragen durch den Standard geklärt werden können (Interviewpartner 7). Darüber hinaus erhofft sich Interviewpartner 10, dass durch den neuen Prüfungsstandard mehr Klarheit bezüglich der Auslegung der Prüfungssicherheit erzeugt werden könne. Zusätzlich solle durch den ISSA 5000 klarer abgegrenzt werden können, welche Prüfungsschwerpunkte und Prüfungshandlungen den jeweiligen Zusicherungsniveaus unterliegen (Interviewpartner 10). Auch Vorgaben für zugelassene dritte Prüfungsdienstleister sollen durch den ISSA 5000 konkretisiert werden (Interviewpartner 12). Ein Gesprächspartner verdeutlichte, dass es einen Prüfstandard bräuchte, welcher die CSRD streng auslege und ein Prüfungsurteil ermögliche, welches zur Zielsetzung der CSRD im geeigneten Ausmaß beiträgt (Interviewpartner 9). Insgesamt erhoffen sich die Interviewteilnehmer mehr Prüfungssicherheit durch das Inkrafttreten des ISSA 5000 (Interviewpartner 7, 10).

2.6.3 Risiken

Inhärente Risiken

Makroökonomische Risiken: Interviewpartner 9 äußerte sich besorgt über die wirtschaftlichen Unsicherheiten, insbesondere in Zeiten von Krisen, die die Prioritäten von Unternehmen verschieben könnten. Unternehmen könnten in solchen Situationen versuchen, die

Berichterstattung über wesentliche Nachhaltigkeitsthemen zu umgehen bzw. weniger priorisieren, da das finanzielle Überleben hier oft in den Vordergrund rücken würde. Die Unsicherheit über die zukünftigen Anpassungen der Regularien wurde mehrfach betont (Interviewpartner 2, 7, 12). Es bestehe die Gefahr, dass Unternehmen durch ungenaue Interpretation der Regularien rechtliche Konsequenzen riskieren (Interviewpartner 4, 12).

Branchenspezifische Risiken

Interviewpartner 5 betont die Unterschiede in der Wesentlichkeit zwischen verschiedenen Branchen. So müssten sich Automobilhersteller bspw. intensiv mit CO2-Emissionen befassen, während bei anderen Branchen wie der Chemieindustrie andere Umweltaspekte im Vordergrund stehen würden. Ein weiterer Befragte gab an, dass technologische Entwicklungen und Veränderungen in der Branche (z. B. Elektromobilität) dazu führen, dass bestimmte Themen plötzlich an Bedeutung gewinnen oder verlieren.

Mandantenspezifische Risiken

Insbesondere komplexe Konzernstrukturen und internationale Standorte würden Schwierigkeiten bereiten, eine konsistente Wesentlichkeitsanalyse durchzuführen (Interviewpartner 1, 5, 6). So könnten kulturelle Unterschiede zwischen Ländern oder Standorten zu unterschiedlichen Schwerpunkten bei der Wesentlichkeit führen (Interviewpartner 9). Interviewpartner 5 hebt hervor, dass es hinsichtlich der Aggregation von Informationen bei großen Mischkonzernen wenig sinnvoll sei, einen einzigen Nachhaltigkeitsbericht für das gesamte Unternehmen zu erstellen, da die Unternehmen im Portfolio unterschiedlich seien.

Prüffeldspezifische Risiken

Ein zentrales Problem bei der Beurteilung der Wesentlichkeitsanalyse ist die mangelnde Qualität und Vollständigkeit der Daten, insbesondere in Bezug auf Lieferketten und Nachhaltigkeitskennzahlen (Interviewpartner 5, 9, 11). Diesbezüglich spricht Interviewpartner 3 auch von der Schwierigkeit alle relevanten Daten aus verschiedenen Geschäftsbereichen und Standorten zusammenzuführen. Ferner fehlen laut Interviewpartner 7 und 9 häufig geeigneten Vergleichsmaßstäbe, um die Relevanz von Nachhaltigkeitsthemen objektiv bewerten zu können. Interviewpartner 12 hebt die Bedeutung der Dokumentation in der Wesentlichkeitsanalyse hervor und weist darauf hin, dass die Dokumentation oft unzureichend ist.

Kontrollrisiken

Interviewpartner 6 führt an, dass es oft keine internen Kontrollmechanismen gibt, die sicherstellen, dass die Wesentlichkeitsanalyse korrekt durchgeführt wird.

Assertions in der Prüfung

Interviewpartner 12 hob die Relevanz des prüferischen Risikos vor dem Hintergrund der davon abhängenden Prüfungsstrategie hervor. Über die Hälfte der Interviewpartner sehen die Vollständigkeit der Wesentlichkeitsanalyse als relevante Assertion, da die größte Gefahr darin bestehe, dass Unternehmen wesentliche Themen auslassen, sei es absichtlich oder durch Datenlücken (2, 5, 6, 7, 10–12). Zudem wurde die Genauigkeit bzw. Richtigkeit als zu adressierende Assertion angeführt (Interviewpartner 5, 7, 9–12). Interviewpartner 7 führte dazu aus, dass bspw. die Matrix in der Wesentlichkeitsanalyse richtig gerechnet sein müsste. Interviewpartner 6 hingegen sieht die Genauigkeit erst als relevante Assertion bei der Prüfung der konkreten Datenpunkten.

Entdeckungsrisiko und Prüfungsrisiko

Viele Prüfer haben noch keine ausreichende Erfahrung mit der Wesentlichkeitsanalyse, was zu Unsicherheiten führe (Interviewpartner 9). Zudem fehle es an standardisierten Prozessen (Interviewpartner 3), um die Wesentlichkeitsanalyse anhand von klaren Kriterien und Maßstäben konkret zu bewerten (Interviewpartner 12).

2.6.4 Prüfungsplanung

Personalplanung

Grundsätzlich schätzten die Gesprächspartner ihre eigenen Fähigkeiten als ausreichend ein, um die Wesentlichkeitsanalyse prüfen zu können. (Interviewpartner 5–7, 10, 12). Für gewisse Nischenthemen innerhalb der Wesentlichkeitsanalyse sahen sie allerdings die Notwendigkeit von zusätzlicher Expertise. Zur Prüfung der Wesentlichkeitsanalyse müssten gegebenenfalls Experten hinzugezogen werden, damit alle fachlichen Themen abgedeckt werden können (Interviewpartner 4–7, 11, 12). Darunter fielen beispielsweise IT-Prüfer, Techniker, Prüfingenieure, Geologen, Umweltwissenschaftler, Statistiker oder Gutachter (Interviewpartner 11, 12). Für die Prüfungsplanung sei deshalb zu berücksichtigen, welche Verantwortlichkeiten und Ressourcen gebraucht und eingeplant werden müssten

(Interviewpartner 7). Für benötigtes Spezialwissen, welches für eine Prüfung der Wesentlichkeitsanalyse gebraucht wird, werden momentan Netzwerke zu diversen Experten aufgebaut (Interviewpartner 6). Teilweise sehen die Interviewpartner eine Herausforderung darin, dem Aufwand der Prüfung personell nachzukommen (Interviewpartner 5, 10). Es seien bereits Überlegungen angestellt worden, einzelne Mandate an andere Wirtschaftsprüfungsgesellschaften abzugeben, da die notwendigen Kapazitäten nicht vorhanden sein werden (Interviewpartner 5). Für die Kanzleien bedeutet der steigende Arbeitsaufwand, dass zusätzliche Mitarbeitende eingestellt werden und die bestehenden Prüfer zeitintensive Fortbildungen durchlaufen müssen (Interviewpartner 4, 5, 12). Die Interviewteilnehmer waren sich in dem Punkt wiederholt einig, dass die Prüfung mindestens so viele personelle Kapazitäten binden würde, wie es die Prüfung der finanziellen Berichterstattung bereits tut (Interviewpartner 1, 4, 5, 9, 10).

Zeitplanung und Berücksichtigung sonstiger Einflussfaktoren bei der Prüfungsdurchführung

Bei der Zeitplanung sind diverse Faktoren zu berücksichtigen, darunter der Entwicklungsfortschritt des jeweiligen Mandanten, der Umfang der Wesentlichkeitsanalyse selbst und das Ausmaß, in welchem der Prüfer bereits bei der Erstellung mit eingebunden wurde (Interviewpartner 1, 12). Der Mandant müsse darüber aufgeklärt werden, dass die Wesentlichkeitsanalyse der Ausgangspunkt für alle weiteren Prüfschritte sei und die Ergebnisse der Prüfung den weiteren Prüfungsablauf bestimme (Interviewpartner 7). Insbesondere für die Prozessaufnahme sollte viel Zeit eingeplant werden (Interviewpartner 9). Regelmäßig werden Wirtschaftsprüfer schon vor Beginn der typischen Prüfungszeit zwischen Oktober und April bei der Erstellung der Wesentlichkeitsanalyse konsultiert, wobei die Kanzleien hier teilweise begleitend tätig werden (Interviewpartner 1, 5). Sofern beide Seiten frühzeitig aufeinander zugehen, können kritische Themen rechtzeitig identifiziert werden oder auch etwaige Prüfungsschwerpunkte kommuniziert werden, sodass Mandanten rechtzeitig Maßnahmen ergreifen können (Interviewpartner 2, 7). Die Wesentlichkeitsanalyse sollte auf der Seite des Erstellers so frühzeitig wie möglich, spätestens ein halbes Jahr vor Abschluss des laufenden Geschäftsjahres abgeschlossen sein, damit eine adäquate Prüfung möglich ist (Interviewpartner 2, 5, 7, 11). Insbesondere für Erstanwender müsse momentan noch viel Zeit für Diskussionen und Beratung eingeplant werden, da manche Unternehmen das Konzept der Wesentlichkeit gem. ESRS nicht korrekt anwenden und – insbesondere im ersten Anwendungsjahr – eine hohe Fehleranfälligkeit erwartet wird (Interviewpartner 1, 5, 8). „Manche Unternehmen haben […] [nicht nachvollziehbare] Vorstellungen mit zwei bis drei wesentlichen Themen und da bedarf

es dann einfach nochmal viel Zeit in die Diskussion zu gehen.“ (Interviewpartner 2). Die Prüfungsplanung, sprich der zeitliche Umfang und die Zeiträume, in welchen einzelne Prüfschritte vorgenommen werden, sollte gemeinsam mit dem Mandanten geplant werden (Interviewpartner 7). Zudem sind interne Abstimmungen im Prüfungsteam zwecks Besprechung der Prüfungsstrategie einzuplanen (Interviewpartner 2). Ferner tangiert das Geschäftsmodell, der Umfang der Lieferkette sowie die Bereitschaft des Mandanten Ressourcen bereitzustellen die Prüfungsplanung (Interviewpartner 1). Auch die Komplexität, der Stand der Digitalisierung sowie Konzernstrukturen und die Geschäftsaktivitäten sind beeinflussende Faktoren (Interviewpartner 2). Bei der zeitlichen Planung ist zudem die Größe des Unternehmens, die Erfahrung des Prüfers und die personelle Stärke des Prüfungsteams zu berücksichtigen (Interviewpartner 5, 6). Die Prüfung der Wesentlichkeitsanalyse nimmt im Verhältnis zur gesamten Prüfung der Nachhaltigkeitsberichterstattung ungefähr ein Drittel der Zeit in Anspruch (Interviewpartner 3, 8, 11). Für die Durchführung planen die Interviewpartner – nach Abhängigkeit der oben erläuterten Einflussfaktoren – wenige Tage über mehrere Wochen bis hin zu drei Monaten ein (Interviewpartner 1, 5, 9). Aufgrund der Neuartigkeit vieler Offenlegungsanforderungen sowie der allgemeinen Aktualität und Dynamik des Themas wird insbesondere die Erstprüfung umfangreicher erwartet (Interviewpartner 2, 6, 7, 9, 11). Dieser Zusatzaufwand marginalisiert sich allerdings in den Folgejahren, da Erstprüfungseffekte wegfallen (Interviewpartner 6, 7, 11).

2.6.5 Prüfungsdurchführung

Orientierung bei der Prüfung

Die Interviewpartner sind sich weitgehend einig, dass eine Standardisierung und klare Prüfmethoden notwendig sind, um die Prüfung von Nachhaltigkeitsberichten effektiv durchzuführen. Interviewpartner 2, 5 und 7 betonen die Bedeutung einer standardisierten Prüfmethodik, die in Teilen bereits existiere und an der sich Prüfungsteams orientieren. Sie sehen in den Standards eine klare Vorgabe für den Prüfungsprozess, von der Planung bis zum Abschluss, wohingegen Interviewpartner 1, 3 und 6 die Notwendigkeit eines flexiblen und pragmatischen Umgangs hervorheben, um den Herausforderungen mit den bisherigen Standards gerecht zu werden. Interviewpartner 2, 5, 7 und 12 machen auf die Bedeutung von Branchenstandards für die Wesentlichkeitsanalyse aufmerksam und unterstreichen die Notwendigkeit des Peer-Vergleichs.

Prüfungsschwerpunkt

Die Wesentlichkeitsanalyse stellt vorerst für alle Interviewpartner einen zentralen Prüfungsschwerpunkt dar, da sie die Grundlage für die gesamte Nachhaltigkeitsberichterstattung bildet. So wird die Analyse bspw. als „Kernstück der Prüfung" bezeichnet (Interviewpartner 6). Es gibt jedoch Unterschiede in der Einschätzung, wie sich dieser Schwerpunkt in Zukunft entwickeln wird. Während mehrere Interviewpartner erwarten, dass die Bedeutung der Wesentlichkeitsanalyse in Zukunft zurückgeht (Interviewpartner 5, 6, 7) sobald sich die Prozesse eingespielt haben, betont Interviewpartner 11, dass die Wesentlichkeit immer ein Prüfungsschwerpunkt bleiben wird. Interviewpartner 7 führt aus, dass die Analyse im Folgejahr oft nur auf Veränderungen überprüft und dann auf bestehenden Ergebnissen aufgebaut wird.

Herausforderungen bei der Prüfung der Wesentlichkeitsanalyse

Die Prüfung der Nachhaltigkeitsberichterstattung bringt eine Reihe von spezifischen Herausforderungen mit sich, die sich auf mehrere Ebenen abspielen. Diese betreffen sowohl die fehlende Standardisierung und Vergleichbarkeit (Interviewpartner 3, 6, 10). So heben Interviewpartner 4 und 6 hervor, dass es an normativen und greifbaren Zahlenwerken – wie klar definierten Wesentlichkeitsschwellen – fehle, auf deren Basis Entscheidungen getroffen werden. „Es gibt immer ein gewisses Ermessen, anders als bei Zahlen, wo es klare Fakten gibt." In diesem Kontext betonen Interviewpartner 7 und 9 die Unsicherheiten bzw. Auslegungsfragen, die bei der Festlegung von Wesentlichkeitsgrenzen bestehen und seitens der Regulierer nicht adressiert seien. Interviewpartner 1 und 6 sehen aufgrund der Interpretationsspielräume und keinen einheitlichen Prüfungsvorgehen die Gefahr, dass die Beurteilung der Wesentlichkeitsanalyse ceteris paribus bei wechselndem Wirtschaftsprüfer zu unterschiedlichen Ergebnissen führen könne, da Sachverhalte heterogen bewertet werden können. Ein gemeinsames Verständnis müsse im ersten Berichtsjahr durch den laufenden Austausch innerhalb des Berufsstandes gefunden werden, sodass Unsicherheiten ausgeräumt werden können (Interviewpartner 2, 9). Zudem berichtet Interviewpartner 3 über branchenabhängige Herausforderungen, die keine universelle Blaupause für die Prüfung der Wesentlichkeitsanalyse zulassen würden. So sieht Interviewpartner 11 insbesondere bei Branchen mit komplexen Lieferketten enorme Schwierigkeiten bei der Überprüfung von externen Daten. Diese Problematik verschärfe sich im Kontext internationaler Lieferketten. Dahingehend berichtet Interviewpartner 5 über Widerstände und Missverständnisse insbesondere in Ländern, in denen die ESG-Berichterstattung weniger verbreitet ist. Die Informationsbeschaffung ist gemäß Interviewpartner 3 das „A und O" jeder Prüfung, wobei

regelmäßig unterschiedliche Systeme und Datensätze im Einsatz sind, die gut koordiniert und aufbereitet werden müssen. Die Wesentlichkeitsanalyse bleibe dynamisch und stark von Stakeholder-Interessen geprägt, bis klare Branchenstandards etabliert seien (Interviewpartner 12). Gleichzeitig wird die Berichterstattung zunehmend anspruchsvoller, da Unternehmen langfristig nicht nur qualitative, sondern auch präzise quantitative Angaben hinsichtlich finanzieller Auswirkungen machen müssen. Hierbei sieht Interviewpartner 9 noch sehr viel Spielraum und Dynamik, die gleichzeitig mit Unsicherheiten in der Erstellung und Prüfung verbunden seien.

Relevante Expertise bei der Prüfungsdurchführung

Die Prüfung der Nachhaltigkeitsberichterstattung erfordert nicht nur eine Erweiterung der traditionellen Wirtschaftsprüfungskompetenzen, sondern auch den intensiven Einsatz von externen und internen Experten (Interviewpartner 4–7, 11, 12). Branchenspezifische Kenntnisse, technische Expertise (Interviewpartner 5) und das fortlaufende Lernen (Interviewpartner 6) seien entscheidend, um den Anforderungen gerecht zu werden. Diese Entwicklungen werden das Berufsbild der Wirtschaftsprüfer nachhaltig verändern, da sie immer mehr auf Spezialisten angewiesen sind und sich kontinuierlich an die neuen Herausforderungen der ESG-Berichterstattung anpassen müssen (Interviewpartner 5).

2.6.6 Prüfungshandlungen

Die Interviewpartner 3 und 7 waren optimistisch, dass sich hinsichtlich der Prüfungshandlungen Best Practices und Leitfäden entwickeln würden, anhand derer ein einheitliches Vorgehen geschaffen werden könne (Interviewpartner 3, 7). Außerdem würden sich in Zukunft Industrie-Benchmarks etablieren, welche zu einer einheitlichen IRO-Bestimmung beitragen könnten (Interviewpartner 7).

Themen identifizieren (Shortlist & Longlist)

In dieser Phase werden potenziell relevante Themen auf eine Longlist gesetzt. Die Identifikation der Themen basiert oft auf einem breiten Spektrum von Datenquellen, internen und externen Analysen sowie Branchenbenchmarks. Die Abgrenzung, Bewertung und Entwicklung der Longlist stellt eine bedeutende Prüfungshandlung dar (Interviewpartner 2, 6, 10). Interviewpartner 9 erläutert, dass bspw. eine „Schatten-Wesentlichkeitsanalyse“ basierend auf Kenntnissen über das Unternehmen oder externen Stakeholder-Informationen erstellt werde, wofür allerdings ein Verständnis zu generischen und branchentypischen IROs vorliegen

müsse. Der Nachvollzug der Erstellung der Wesentlichkeitsanalyse hat laut Interviewpartner 7 eine zentrale Bedeutung bei der Prüfung. Daher finde hierzu eine umfangreiche Prozessaufnahme (Interviewpartner 3, 9) im Rahmen von vielen Gesprächen mit den jeweiligen Verantwortlichen in den Fachabteilungen statt (Interviewpartner 1, 2, 3). Erkenntnisse aus anderen Berichten (Interviewpartner 5) sowie Peer- oder Medienanalysen seien bei der Beurteilung der Vollständigkeit hilfreich (Interviewpartner 7). Interviewpartner 9 führt an, dass in seiner Abteilung in Zusammenarbeit mit unterschiedlichen Stakeholdern – wie bspw. Umweltverbänden – eine umfangreiche Datensammlung durchgeführt wurde, wobei Erkenntnisse zu Auswirkungen, Risiken und Chancen in diversen Branchen zusammengefasst wurden. Diese Sammlung hilft ihm zur Folge, relevante Erwartungen hinsichtlich der IROs zu definieren. Ein weiteres Instrument seien Checklisten, mit denen der Prüfer die wesentlichen Themen der Longlist abstimmen könne (Interviewpartner 7). Ferner kämen als typische Prüfungshandlungen Befragungen Dritter in Frage (Interviewpartner 11, 12). Es gäbe häufig Meinungsverschiedenheiten bezüglich Themen, die in einer Wesentlichkeitsanalyse identifiziert werden müssten. In so einem Fall verweist Interviewpartner 1 auf prüferisches Ermessen und die Begründung des Mandanten. Solange diese für den Prüfer plausibel und nachvollziehbar ist, ließe sich kein Widerspruch einwenden.

Die Reduktion von der Longlist zur Shortlist sei ein notwendiger Schritt, um sicherzustellen, dass relevante Inhalte im Nachhaltigkeitsbericht dargestellt werden (Interviewpartner 6). Die Prüfmethodik umfasse in dieser Phase eine Bewertung anhand der Anforderungen, die für die identifizierten Datenpunkte gelten, sowie deren Wesentlichkeitseinschätzung. Mithilfe von Stichproben könne man dann tiefer in die Thematik einsteigen, um die Richtigkeit zu adressieren (Interviewpartner 2). Dabei werden insbesondere Nachweise, wie die Wesentlichkeitsmatrix, einbezogen (Interviewpartner 7).

Überprüfung der Einbindung von Stakeholdern

Interviewteilnehmer 1 und 11 betonten, dass das Wissen über das Unternehmen half, relevante Stakeholder zu identifizieren, die in den Entstehungsprozess der Wesentlichkeitsanalyse eingebunden werden sollten. Dies könnte durch Fragebögen oder E-Mail-Kontakt geschehen, wobei die Dokumentation dieser Einbindung der Prüfung zugänglich gemacht werden sollte (Interviewpartner 6). Grundsätzlich sahen es Interviewteilnehmer auch als Möglichkeit, ein Gespräch mit den Stakeholdern zu führen, um die Einbindung beurteilen zu können (Interviewpartner 11), wobei die Antwortquote nicht unbedingt hoch sei (Interviewpartner 6).

Daher könnte auch mithilfe von Regressions- oder Korrelationsanalysen zu Lieferanten, Kunden oder anderen Vertragsbeziehungen geprüft werden, ob alle wesentlichen Kreditoren, Lieferanten oder auch Kunden angesprochen worden sind (Interviewpartner 11).

Prüfungshandlung hinsichtlich der Lieferkette

Die Einbindung der Lieferkette sei für viele Mandanten eine herausfordernde Aufgabe (Interviewpartner 7), wobei einige Unternehmen die Bedeutung ihrer vor- und nachgelagerten Lieferkette angesichts der Prüfung bagatellisieren würden. Der Vergleich innerhalb ähnlicher Branchen sowie ein Austausch mit Experten zu kritischen Themen innerhalb einer Lieferkette könnten den Prüfer dabei unterstützen, eine Erwartungshaltung zu bilden. Zur Überprüfung müssten Informationsquellen, Kontakte und entsprechende Nachweise offengelegt und gewürdigt werden. Relevant sei dabei vor allem die Konsistenz und Transparenz auf Seiten des Mandanten (Interviewpartner 6). Interviewpartner 12 sah eine besondere Herausforderung bei Rohstoffen, die an Spotmärkten beschafft würden. In einem solchen Fall könnte es Schwierigkeiten geben, zu identifizieren, welcher Lieferant den Rohstoff produziere (Interviewpartner 12). Ein anderer Interviewpartner stellte in Frage, ob für die Einbindung der Lieferkette Schätzwerte genutzt werden können. Diese können in der Prüfung gewürdigt werden, solange die Annahmen transparent nachvollzogen werden können (Interviewpartner 1).

Kriterien Quellen und Datenqualität

Der Prüfer sollte die Herkunft der Daten und Quellen prüfen (Interviewpartner 8, 10). Die Wesentlichkeitsschwellen seien sehr individuell und es gäbe noch kein einheitliches Vorgehen, wie die Kriterien und Quellen überprüft werden können (Interviewpartner 1, 2, 6). Die Methodik und Anwendung hinter den Kriterien müsse verständlich für einen Prüfer sein. Auch hier gelte der Grundsatz der Plausibilität für die Prüferseite. Solange das berichtende Unternehmen nachvollziehbar verdeutlichen kann, welche Quellen und Kriterien innerhalb der Wesentlichkeitsanalyse angewendet wurden, gäbe es keinen Anlass für einen Widerspruch (Interviewpartner 3, 6, 7). Benchmark-Analysen und das Verplausibilisieren von Schätzungen seien relevante Prüfungshandlungen innerhalb der Bewertung von selbst gewählten Quellen und Kriterien (Interviewpartner 11). Für die Feststellung der Wesentlichkeit käme eine Einschätzung mittels einer möglichst entscheidungsnützlichen Skala in Frage (Interviewpartner 1). Dahingehende Vorgaben seitens der EFRAG seien für die Überprüfung hilfreich und sollten

berücksichtigt werden. Solange die Datenermittlung validiert werden könne, seien auch die Prozesse der Wesentlichkeitsanalyse als verlässlicher einzuschätzen (Interviewpartner 2).

2.6.7 Prüfungssicherheit

Es gibt unterschiedliche Ansichten darüber, wie stark sich die beiden Sicherheitsniveaus im ersten Jahr voneinander unterscheiden. Einige Experten betonen, dass es im ersten Jahr hinsichtlich der Prüfung der Wesentlichkeitsanalyse kaum Unterschiede zwischen den beiden Sicherheitsstufen geben wird (Interviewpartner 1, 5). Da die Wesentlichkeitsanalyse einen Prüfungsschwerpunkt darstelle, müsse die Prüfung mit einem hohen Detailierungsgrad durchgeführt werden (Interviewpartner 6). Eine genauere Differenzierung zwischen den beiden Zusicherungsniveaus erhoffen sich die Interviewteilnehmer durch eine Finalisierung des Prüfungsstandards und zukünftige Klarstellungen des IDWs (Interviewpartner 3, 6). Der geplante Übergangszeitraum zwischen den beiden Prüfurteilen sei angedacht, um die Prüfung weiterzuentwickeln und Erfahrungswerte zu sammeln, mit denen sich eine Grenze ziehen ließe (Interviewpartner 12). Solange diese Grenze nicht vorhanden sei, sollte besondere Vorsicht geboten und das mögliche Risiko bedacht werden (Interviewpartner 9).

Begrenzte Sicherheit

Der Fokus einer Prüfung mit begrenzter Sicherheit liege auf dem Verplausibilisieren (Interviewpartner 11). Die Wesentlichkeitsanalyse müsse demnach mit den Erwartungen der Prüfer übereinkommen (Interviewpartner 9, 10), wozu Gespräche, Dokumentensichtung und der Abgleich mit den entsprechenden Standards und Erwartungswerten der Prüfer (Interviewpartner 3) mittels analytischer Prüfungshandlungen wie der Nutzung von Trend- oder Abweichungsanalysen genüge (Interviewpartner 8, 10, 12). Ein prozentuales Sicherheitsniveau für eine begrenzte Sicherheit gaben die Gesprächspartner nicht an. Interviewpartner 10 sah hier die Schwierigkeit gegeben, dass eine entsprechend niedrige Quantifizierung des begrenzten Sicherheitsgefühls nicht möglich sei.

Hinreichende Sicherheit

Die Experten sind sich einig, dass die Abgrenzung zwischen begrenzter und hinreichender Sicherheit vor allem durch den Dokumentationsaufwand, die Prüfungstiefe und den Einsatz analytischer Prüfungen bestimmt wird. Während die begrenzte Sicherheit als pragmatischer Ansatz für den Einstieg gesehen wird, erfordert die hinreichende Sicherheit umfassendere Einzelfallprüfungen (Interviewpartner 1, 3, 4, 10, 11). Zukünftig seien Aufbau- und

Funktionsprüfungen von internen Kontrollsystemen notwendig, um das angestrebte Sicherheitsniveau – von mindestens 95 % (Interviewpartner 4, 8, 12) – zu erreichen (Interviewpartner 2, 3, 7, 11). Aktuell seien die Kontrollsysteme und IT-Systeme auf Mandantenseite noch nicht ausreichend, um einen derartigen Fokus in der Prüfung legen zu können (Interviewpartner 7). Um dieses Sicherheitsniveau innerhalb einer Prüfung zu erreichen, müssten die Prüfungshandlungen detaillierter und umfassender gestaltet werden (Interviewpartner 4, 8).

2.6.8 Bestätigungsvermerk

Inhalt des Prüfungsurteils:

Eine explizite Ausgestaltung des Vermerks hätte sich noch nicht herauskristallisiert (Interviewpartner 2, 4, 7), wobei erwartet wird, dass sich ähnlich wie in der Finanzberichterstattung ein Berufsstandard entwickelt, sodass die Testate in Zukunft standardisiert und einheitlich aussehen werden (Interviewpartner 1). Interviewpartner 11 erwartet, dass die Prüfvermerke, welche erstmals im Jahr 2025 veröffentlicht werden, maßgeblich als Orientierung dienen. Die Wesentlichkeitsanalyse sei ein bedeutender Bestandteil des Vermerks, sollte jedoch nicht zu detailliert dargestellt werden (Interviewpartner 1, 4). Der Vermerk werde sich eher auf Informationen konzentrieren, die für den Markt relevant sind, und weniger auf die Wesentlichkeitsanalyse selbst (Interviewpartner 6). Die Dokumentation der Prüfungsschritte bleibe zentral und werde umfassend durchgeführt, inklusive der Begründung von Entscheidungen und Prozessen (Interviewpartner 8, 9).

Bedeutung des Prüfungsurteils und etwaige Einschränkungen:

Interviewpartner 1, 7 und 10 rechnen mit einer größeren Anzahl an eingeschränkten Testaten, unter anderem aufgrund fehlerhafter Wesentlichkeitsanalysen (Interviewpartner 12). Im Austausch mit der BaFin sowie dem IDW sei diese Erwartungshaltung bestätigt worden (Interviewpartner 10). Drei der Interviewteilnehmer verdeutlichen diesbezüglich nochmals die Vorteile, welche eine frühzeitige Zusammenarbeit zwischen Prüfer und Mandant mit sich bringe (Interviewpartner 1, 2, 7).

Sei die Bedeutung in den ersten Jahren noch verhältnismäßig gering (Interviewpartner 4, 11), so nehme diese mit der Zeit stark zu. Drei Experten sehen den Prüfvermerk als Hebel für Veränderungen und erwarteten insbesondere für jüngere Stakeholder und Banken eine wachsende Relevanz (Interviewpartner 1, 2, 10). Die Bedeutung des Prüfurteils würde steigen,

wenn die Nichteinhaltung von Nachhaltigkeitskennzahlen zu rechtlichen Sanktionen oder Marktausschluss führen würde (Interviewpartner 3). Einigkeit herrschte darüber, dass das Urteil zur Wesentlichkeitsanalyse und etwaige Einschränkung im ISAE 3000 Vermerk keinen Einfluss auf andere Urteile wie jenes zur Unternehmensfortführung bzw. den finanziellen Bestätigungsvermerk haben werde (Interviewpartner 2, 3, 5, 11). Allerdings gäbe es durch den ausgeweiteten Prognosezeitraum in der Nachhaltigkeitsberichterstattung einen verzögerten Zusammenhang, da langfristige Risiken, die durch die doppelte Wesentlichkeit identifiziert werden, sich erst später in der finanziellen Berichterstattung wiederfinden (Interviewpartner 12).

2.7 Schlussbetrachtung

2.7.1 Diskussion der Ergebnisse

Mittels Auswertung der durchgeführten Interviews wurde ein umfassender Einblick in die Durchführung der Prüfung der doppelten Wesentlichkeitsanalyse ermöglicht. Ferner wurden damit einhergehende Chancen und Herausforderungen ausgelotet. Neben der aktuellen Lage in den Wirtschaftsprüfungskanzleien konnten auch Erkenntnisse zum Stand der Umsetzung bei den berichtspflichtigen Unternehmen gewonnen werden.

Einordnung der Ergebnisse vor dem Hintergrund einschlägiger wissenschaftlicher Theorien

Die Erkenntnisse stehen im Kontext etablierter Theorien wie der Prinzipal-Agent-Theorie und der Theorie der Entscheidungsnützlichkeit und bieten eine geeignete Grundlage, um die Rolle der Prüfung und deren Auswirkungen auf die Stakeholder zu beleuchten.

Es zeigt sich, dass die Prüfung der Wesentlichkeit als entscheidendes Instrument gesehen wird, um sowohl die Qualität der Berichterstattung als auch die Glaubwürdigkeit gegenüber Stakeholdern zu erhöhen. Einigkeit herrscht darüber, dass die Prüfung insbesondere die Sicherstellung der Vollständigkeit aller relevanten Auswirkungen, Chancen und Risiken hinsichtlich beider Materialitätsperspektiven zum Ziel hat (Interviewpartner 2, 5, 6, 7, 10–12). So kann die momentan herrschende Problematik begegnet werden, dass selektiv positive Nachhaltigkeitsaspekte berichtet werden (Seele/ Gatti 2017, S. 248). Interviewpartner 5, 6 und 11 wiesen darauf hin, dass ohne eine Prüfung wesentliche Themen entweder unterschlagen oder der Bericht durch zu viele unwesentliche Themen verwässert werden könnte. Diese Überladung von Berichten mit irrelevanten Themen könnte die Entscheidungsfindung für Stakeholder erschweren und die Entscheidungsnützlichkeit der Berichte mindern. Dies kann zum einen im

Kontext der Theorie der Entscheidungsnützlichkeit betrachtet werden und adressiert zugleich das *moral hazard* Problem im Kontext der Prinzipal-Agent Theorie. Die zunehmende Bedeutung von leistungsabhängigen Vergütungssystemen auf Basis von Nachhaltigkeitskennzahlen (Interviewpartner 5) kann zu Interessenkonflikten und Manipulationen führen, wenn das Management primär auf die Maximierung der eigenen Vergütung fokussiert ist. Die Aufgabe der Prüfer besteht darin diese Informationsasymmetrien zu überwinden, indem sie eine objektive und verlässliche Prüfung der Nachhaltigkeitskennzahlen durchführen und sicherstellen, dass die Berichterstattung den tatsächlichen Risiken und Chancen des Unternehmens entspricht. Dies ist entscheidend, um das Vertrauen der Investoren und anderer Stakeholder in die Nachhaltigkeitsberichterstattung zu stärken und das Risiko von *moral hazard* zu minimieren.

Andererseits wurden oft die Herausforderung hinsichtlich der Datenqualität angesprochen. Interviewpartner 9 hob die Schwierigkeit bei Mischkonzernen alle relevanten Daten aus verschiedenen Unternehmensbereichen und Standorten zu aggregieren hervor. Dies erschwert die Durchführung einer konsistenten Wesentlichkeitsanalyse und kann die Entscheidungsnützlichkeit der Information reduzieren.

Status quo in der Wirtschaftsprüfung und die Vorbehaltsaufgabe

Ungefähr die Hälfte der Interviewpartner sprachen sich für die Vorbehaltsaufgabe aus, wobei sie die Vorteile des Berufsstands und seine Erfahrung mit komplexen Prüfungen betonten. Gleichzeitig wurde darauf verwiesen, dass spezifische technische Expertise, wie im Bereich von Umweltgutachten, weiterhin von externen Spezialisten wie dem TÜV oder Umweltingenieuren beigesteuert und Expertennetzwerke aufgebaut werden sollten. Mehrere Interviewpartner sahen den Abschlussprüfer als prädestiniert für die effiziente Prüfung des Nachhaltigkeitsberichts, da so eine Entkopplung von Finanz- und Nachhaltigkeitsinformationen vorgebeugt werden kann und Synergien gehoben werden können (so auch Quick 2024b, S. 468; Velte/Stawinoga 2022, S. 1143). Allerdings wurde die Möglichkeit der Bestellung anderer Wirtschaftsprüfer teilweise, aus pragmatischen Gründen der sonstigen Überforderung, auch begrüßt. Dies steht im Zusammenhang mit dem Fachkräftemangel in den Kanzleien und Unternehmen, der die erfolgreiche Durchführung von Prüfungen behindert. Der zunehmende Bedarf an spezialisierten Fachkräften, die sowohl die technischen als auch die regulatorischen Anforderungen der Nachhaltigkeitsberichterstattung verstehen, übersteigt aktuell das verfügbare Angebot an qualifiziertem Personal. Dieser Mangel

ist besonders für kleine Unternehmen problematisch, da diese Schwierigkeiten haben, entsprechende Ressourcen kurzfristig aufzubauen. Perspektivisch wird das neue Aufgabengebiet allerdings als Chance gesehen, das Nachwuchsproblem im Berufsstand zu adressieren (Velte 2023, S. 1297).

Ablauf und Herausforderungen bei der Prüfung

Die Interviews verdeutlichen, dass die Prüfung der doppelten Wesentlichkeitsanalyse ein mehrstufiger und komplexer Prozess ist, der tiefgehende Kenntnisse über das zu prüfende Unternehmen und dessen Branche erfordert (so auch Bozzolan/Miihkinen 2021 im Kontext von Risikoberichten). Insgesamt zeigt sich, dass die Datenqualität und die Auswahl der Quellen und Kriterien entscheidend für die Prüfung der Wesentlichkeitsanalyse sind, gleichzeitig aber auch die größten Herausforderungen darstellen. Die Beschaffung vollständiger und verlässlicher Daten, insbesondere in Bezug auf die Lieferkette, scheitert oft an der Dokumentation seitens der Unternehmen. Eine hohe Unsicherheit und fehlende Standardisierung erschweren zudem die Bewertung dieser Elemente, was die Notwendigkeit für klar definierte Standards und eine bessere Dateninfrastruktur unterstreicht.

Die befragten Wirtschaftsprüfer hoben hervor, dass die Prüfung in den ersten Jahren besonders anspruchsvoll ist, da standardisierte Prüfprozesse noch nicht etabliert sind und wesentliche Unsicherheiten bei der Interpretation der Anforderungen bestehen, wobei der Entwicklungsfortschritt des Mandanten und der Zeitpunkt, zu dem der Prüfer in die Wesentlichkeitsanalyse eingebunden wird, die Prüfung maßgeblich tangieren. Frühe Abstimmungen ermöglichen die rechtzeitige Identifikation kritischer Themen und eine adäquate Prüfungsdurchführung. Erste Erfahrungen sprechen für eine frühestmögliche Einbindung des Prüfers, da so noch Gegenmaßnahmen ergriffen werden können.

Ein zentrales Element der Prüfung ist die Reduktion von einer Longlist potenziell wesentlicher Themen auf eine Shortlist berichtsrelevanter Themen, wobei die Nachvollziehbarkeit des Prozesses eine zentrale Rolle spielt. Die Identifikation potenziell relevanter Themen erfolgt oft auf Basis einer breiten Datenbasis, die sowohl interne als auch externe Quellen umfasst. So wurden bspw. in Ermangelung einschlägiger Branchenstandards zusammen mit Experten eigens hierfür erhobene Datenbasis geschaffen, die Auswirkungen, Risiken und Chancen in diversen Branchen umfasst, um diese mit den Listen der jeweiligen Unternehmen abzugleichen. Prüfer stützen sich auf Checklisten, Befragungen und Stichproben, um die Wesentlichkeit der Themen zu bewerten und die Richtigkeit der Daten zu überprüfen, wobei darauf hingewiesen

wurde, dass es oft an geeigneten Vergleichsmaßstäben fehlt, um die Relevanz von Nachhaltigkeitsthemen objektiv bewerten zu können. Die Experten betonten die Notwendigkeit klarer Standards und Leitlinien, um die Prüfung der Wesentlichkeitsanalyse zu vereinheitlichen. Der ISSA 5000 wird als zukünftiger Prüfstandard eine wichtige Rolle spielen, insbesondere zur Schaffung von Klarheit bezüglich der Prüfungssicherheit und der Abgrenzung der Prüfungshandlungen. Zudem wurde deutlich, dass die Aus- und Weiterbildung der Prüfer entscheidend für die erfolgreiche Bewältigung dieser neuen Prüfungsanforderungen ist. Die traditionellen Prüfungsansätze müssen erweitert werden, um Umwelt und Sozialaspekte adäquat bewerten zu können.

Es gibt Unterschiede zwischen den beiden Zusicherungsniveaus – begrenzte Sicherheit und hinreichende Sicherheit. Die Prüfungshandlungen unterscheiden sich in ihrer Quantität und Tiefe. Disser und Skoluda (2024) bieten hierzu explizite Einblicke entlang des ED ISSA 5000. Die Interviews zeigen, dass die Gewährleistung hinreichender Sicherheit eine Herausforderung darstellt. Insbesondere die Unsicherheiten bei der Auslegung der Wesentlichkeitsgrenzen scheinen ein zentrales Problem darzustellen. Da es noch keine einheitlichen und klar definierten Kriterien gibt, wie Wesentlichkeit objektiv gemessen wird, besteht die Gefahr, dass verschiedene Prüfer zu unterschiedlichen Ergebnissen kommen. Zusammenfassend lässt sich sagen, dass der Hauptunterschied zwischen den beiden Zusicherungsniveaus in der Prüfungstiefe und dem Dokumentationsaufwand liegt. Während die begrenzte Sicherheit auf plausibilitätsorientierten, analytischen Verfahren basiert, erfordert hinreichende Sicherheit deutlich umfassendere Einzelfallprüfungen sowie Aufbau- und Funktionsprüfungen der internen Kontrollsysteme. Aktuell sind die Kontrollsysteme und IT-Systeme auf Mandantenseite noch nicht ausreichend, um einen derartigen Fokus in der Prüfung legen zu können. Quick (2024, S. 468) stellt ebenfalls in Frage, ob eine Prüfung mit hinreichender Sicherheit überhaupt möglich ist.

Der Prüfungsvermerk sollte sich auf Informationen konzentrieren, die für den Markt relevant sind. Hierbei spielt eine transparente und begründete Darlegung des Prüfungsvorgehen seine zentrale Rolle (so auch Velte/Hellriegel 2024, S. 911).

2.7.2 Limitation und Implikation für Forschung und Praxis

Qualitative Forschungsansätze haben ihre Grenzen, die auch in dieser Forschungsarbeit berücksichtigt werden müssen. Ein wesentlicher Punkt ist, dass in qualitativen Studien keine

statistische Generalisierbarkeit abgeleitet werden kann. Zudem basieren die Erkenntnisse zur Vorbehaltsaufgabe auf einer Selbsteinschätzung.

Mit der Verabschiedung des ISSA 5000 ist man einer Forderung der Interviewpartner bereits nachgekommen. So können einheitliche Prüfungsstandards bspw. das Vertrauen von Investoren stärken, welche eine der Hauptzielgruppen der Berichterstattung darstellt (GarciaSanchez et al. 2022, S. 103; RL 2022/2464) ErwG 9). Weitere Präzisierungen und Praxishinweise – wie bspw. von den Interviewpartnern explizit genannten Institut der Wirtschaftsprüfer (IDW) – könnten Unsicherheiten und unterschiedliche Auslegungen in der Prüfungspraxis begegnen. Die Relevanz von Branchenstandards, Leitfäden und vergleichbaren Indikatoren für die Bestimmung von Wesentlichkeitsschwellen ist in den Interviews deutlich hervorgehoben worden. Sie können Prüfern Orientierung liefern, um die Vollständigkeit der Wesentlichkeitsanalyse zu beurteilen und sicherzustellen, dass den Nutzern vergleichbare Nachhaltigkeitsinformationen bereitgestellt werden.

Unternehmen müssen ihre internen Kontrollmechanismen im Hinblick auf die Nachhaltigkeitsberichterstattung und die Wesentlichkeitsanalyse stärken. Die Interviewpartner betonten, dass oft keine ausreichenden Kontrollsysteme vorhanden sind, um sicherzustellen, dass wesentliche Themen korrekt identifiziert und berichtet werden. Es müssen geeignete Strukturen geschaffen werden, um die Datenqualität zu verbessern und eine konsistente und vollständige Wesentlichkeitsanalyse zu gewährleisten. Dies wird insbesondere in Hinblick auf die Prüfung mit hinreichender Sicherheit eine zentrale Rolle spielen (so auch Velte/Stawinoga 2022, S. 1140).

Die Experten erwarten angesichts des bisherigen Umsetzungsstands bei den Unternehmen, dass die Anzahl der eingeschränkten Prüfvermerke nicht unerheblich sein wird. Diese Entwicklung bleibt abzuwarten und ist bei Eintritt auf ihre Wirkung zu untersuchen. Ferner könnten Faktoren untersucht werden, die zu Einschränkungen führen. Für eine dafür geeignete Datengrundlage ist wahrscheinlich aber mindestens die zweite Berichtswelle abzuwarten, da die großen Unternehmen von öffentlichem Interesse aufgrund ihrer Erfahrung mit der Erstellung und freiwilligen Prüfung der Nachhaltigkeitsberichterstattung vermutlich noch nicht von Einschränkungen betroffen sein werden.

Summary

In light of the growing demand for reliable sustainability information, the Corporate Sustainability Reporting Directive (CSRD) introduces expanded reporting obligations as well

as an external audit requirement for a significantly larger user group. The revised concept of double materiality analysis plays a key role in determining the relevant reporting content. Based on 12 expert interviews with auditors, this paper contributes to a better understanding of the required audit processes. Auditing the double materiality analysis requires an in-depth engagement with industry-specific risks and a comprehensive analysis of the underlying data. While limited assurance audits primarily rely on plausibility assumptions and analytical procedures, reasonable assurance requires detailed case-by-case examinations and a thorough analysis of internal control systems. One of the greatest challenges is ensuring data quality and overcoming the lack of guidance. There is a clear need for further standardization, particularly through the development of industry-specific indicators to determine materiality thresholds, as well as targeted capacity building and professional development for auditors. These steps are necessary to meet the increasing audit requirements and to ensure the reliability of the disclosed information for report recipients.

Literaturverzeichnis

Albers, C., & Günther, T. (2010). Disclose or not disclose: determinants of social reporting for STOXX Europe 600 firms. Zeitschrift für Planung & Unternehmenssteuerung, 21(3), 323–347. https://doi.org/10.1007/s00187-010-0113-4.

Ballou, B., Chen, P., Grenier, J. H., & Heitger, D. L. (2018). Corporate social responsibility assurance and reporting quality: Evidence from restatements. Journal Of Accounting And Public Policy, 37(2), 167–188. https://doi.org/10.1016/j.jaccpub pol.2018.02.001.

Baumüller, J. (2024). ESRS-konforme Implementierung der Wesentlichkeitsanalyse: Ausgewählte Fallstricke aus der bisherigen Praxis der neuen Nachhaltigkeitsberichterstattung. Zeitschrift für Corporate Governance, 19(2), 78-83. https://doi.org/ 10.37307/j.1868-7792.2024.02.09.

Baumüller, J., & Sopp, K. (2022). Double materiality and the shift from non-financial to European sustainability reporting: review, outlook and implications. Journal of Applied Accounting Research, 23(1), 8–28. https://doi.org/10.1108/JAAR-04-2021-0114.

Boiral, O., & Heras-Saizarbitoria, I. (2020). Sustainability reporting assurance: Creating stakeholder accountability through hyperreality? Journal of Cleaner Pro-duction, 243, 1–17. https://doi.org/10.1016/j.jclepro.2019.118596.

Bozzolan, S., & Miihkinen, A. (2021). The quality of mandatory non-financial (risk) disclosures: The moderating role of audit firm and partner characteristics. The International Journal of Accounting, 56(02), 2150008. https://doi.org/10.1142/S1094406021500086.

Bundesministerium der Justiz (2024). Gesetz zur Umsetzung der EU-Richtlinie zur Nachhaltigkeitsberichterstattung (CSRD) Informationspapier. https://www.bmjv.de/SharedDocs/Downloads/DE/Gesetzgebung/Dokumente/Infopapier_CSRD_UG.pdf?__blob=publicationFile&v=2.

Bundesregierung (2024). Entwurf eines Gesetzes zur Umsetzung der Richtlinie (EU) 2022/2464 des Europäischen Parlaments und des Rates vom 14. Dezember 2022 zur Änderung der Verordnung (EU) Nr. 537/2014 und der Richtlinien 2004/109/EG, 2006/43/EG und 2013/34/EU hinsichtlich der Nachhaltigkeitsberichterstattung von Unternehmen.

Buschmann, M. (2024). Rede des Bundesministers der Justiz, Dr. Marco Buschmann, anlässlich der Eröffnungsveranstaltung des 74. Deutschen Juristentags in Stuttgart am

25. September 2024. URL: https://www.bmj.de/SharedDocs/Reden/DE/2024/0926_74_Juristentag.html?nn=184772.

Chung, R., Bayne, L., & Birt, J. L. (2024). Determinants of ESG disclosure among listed firms under voluntary and mandatory ESG disclosure regimes in Hong Kong. Journal of Applied Accounting Research, 25(4), 812–836. https://doi.org/10.1108/ JAAR-07-2022-0179.

Clarkson, P., Li, Y., Richardson, G., & Tsang, A. (2019). Causes and consequences of voluntary assurance of CSR reports: International evidence involving Dow Jones

Sustainability Index Inclusion and Firm Valuation. Accounting, Auditing & Accountability Journal, 32(8), 2451-2474. https://doi.org/10.1108/AAAJ-03-2018-3424.

Chung, R., Bayne, L., & Birt, J. (2024). Stakeholder responses to mandating environmental, social and governance reporting in Hong Kong. China Journal of Accounting Research, 17(2), 100353. https://doi.org/10.1016/j.cjar.2024.100353.

Cohen, J. R., & Simnett, R. (2015). CSR and Assurance Services: A Research Agenda. Auditing A Journal Of Practice & Theory, 34(1), 59–74. https://doi.org/10.2308/ajpt-50876

Cuadrado-Ballesteros, B., Martinez-Ferrero, J., & Garcia-Sanchez, I. M. (2017). Mitigating information asymmetry through sustainability assurance: The role of accountants and levels of assurance. International Business Review, 26(6), 1141–1156. https://doi.org/10.1016/j.ibusrev.2017.04.009.

Cui, J., Jo, H., & Na, H. (2018). Does corporate social responsibility affect information asymmetry? Journal of business ethics, 148, 549–572. https://doi.org/10.1007/ s10551-015-3003-8.

DeFond, M., & Zhang, J. (2014). A review of archival auditing research. Journal of accounting and economics, 58(2-3), 275–326. https://doi.org/10.1016/j.jacce-co.2014.09.002.

Deutsches Rechnungslegungs Standards Committee e. V. (2023). CSRD/ESRS: Verschiebung bestimmter ESRS; DRSC Briefing Paper zur CSRD für Drittstaatenunternehmen. https://www.drsc.de/news/verschiebung-bestimmter-esrs-drsc-briefing-pa per-zur-csrd-fuer-drittstaatenunternehmen/.

Deutsches Rechnungslegungs Standards Committee e. V., Accounting Standards Committee of Germany & DRSC. (2021). CSR-STUDIE: Abschlussbericht zur vom BMJV beauftragten Horizontalstudie sowie zu Handlungsempfehlungen für die Überarbeitung der CSR-Richtlinie. https://www.drsc.de/app/uploads/2021/06/210128_ CSR-Studie_final.pdf.

Disser, I., & Skoluda, S. (2024). Entwurf des ISSA 5000: Prüfung der Nachhaltigkeitsberichterstattung mit hinreichender bzw. begrenzter Sicherheit. WPg – Die Wirtschaftsprüfung, 2, 64–73.

Disser, I., & Vogl, S. (2023). Prüfung der Nachhaltigkeitsberichterstattung mit hinreichender Sicherheit. WPg – Die Wirtschaftsprüfung, 5, S. 217–224.

Del Gesso, C., & Lodhi, R. N. (2025). Theories underlying environmental, social and governance (ESG) disclosure: a systematic review of accounting studies. Journal of Accounting Literature. https://doi.org/10.1108/JAL-08-2023-0143.

ED ISSA 5000 (2023). Proposed International Standard on Sustainability Assurance 5000. General Requirements for Sustainability Assurance Engagements and Pro-posed Conforming and Consequential Amendments to Other IAASB Standards von August 2023. IAASB (Hrsg.), International Standard on Sustainability Assurance, Exposure Draft.

EFRAG (2024). EFRAG IG 1: Materiality Assessment Implementation Guidance. https://www.efrag.org/sites/default/files/sites/webpublishing/SiteAssets/IG%201 %20Materiality%20Assessment_final.pdf.

Europäische Kommission (2001). Grünbuch Europäische Rahmenbedingungen für die soziale Verantwortung der Unternehme. https://eur-lex.europa.eu/legal-content/ DE/TXT/PDF/?uri=CELEX:52001DC0366&from=DE.

Europäische Kommission (2023): Delegierte Verordnung (EU) 2023/2772 der Kommission vom 31. Juli 2023 zur Ergänzung der Richtlinie 2013/34/EU des Europäischen Parlaments und des Rates durch Standards für die Nachhaltigkeitsberichterstattung.

Europäische Kommission (2022): Richtlinie (EU) 2022/2464 des Europäischen Parlaments und des Rates vom 14. Dezember 2022 zur Änderung der Verordnung (EU) Nr. 537/2014 und der Richtlinien 2004/109/EG, 2006/43/EG und 2013/34/EU hinsichtlich der Nachhaltigkeitsberichterstattung von Unternehmen.

Europäische Kommission (2019) Mitteilung der Kommission an das europäische Parlament, den europäischen Rat, den europäischen Wirtschafts- und Sozialausschuss und den Ausschuss der Regionen: Der europäische Grüne Deal (2019).

Europäische Kommission (2011) Mitteilung der Kommission an das Europäische Parlament, den Rat, den Europäischen Wirtschafts- und Sozialausschuss und den Ausschuss der Regionen. Eine neue EU-Strategie (2011-14) für die soziale Verantwortung der Unternehmen (CSR).

Europäisches Parlament, Rat der Europäischen Union (2022). Richtlinie (EU) 2022/ 2464 des Europäischen Parlaments und des Rates vom 14. Dezember 2022 zur Änderung der Verordnung (EU) Nr. 537/2014 und der Richtlinien 2004/109/EG, 2006/43/ EG und 2013/34/EU hinsichtlich der Nachhaltigkeitsberichterstattung von Unternehmen.

Fiandrino, S., Di Gromis Trana, M., Tonelli, A., & Lucchese, A. (2022). The multifaceted dimensions for the disclosure quality of non-financial information in revising directive 2014/95/EU. Journal of Applied Accounting Research, 23(1), 274–300. https://doi.org/10.1108/JAAR-04-2021-0118

Fiandrino, S., Tonelli, A., & Devalle, A. (2022). Sustainability materiality research: a systematic literature review of methods, theories and academic themes. Qualitative Research in Accounting & Management, 19(5), 665–695. https://doi.org/10.1108/ qram-07-2021-0141.

Fiechter, P., Hitz, J., & Lehmann, N. (2022). Real Effects of a Widespread CSR Reporting Mandate: Evidence from the European Union's CSR Directive. Journal Of Accounting Research, 60(4), 1499–1549. https://doi.org/10.1111/1475-679x.12424.

Freiberg, J., Terko, S. & Hirnkow, M. (2023): § 17 Prüfung von Nachhaltigkeitsinfor-mationen, in: Freiberg, J., & Lanferman, G. (Eds.). (2023). Haufe ESRS-Kommentar 1. Auflage: Kommentar zu den European Sustainability Reporting Standards. Haufe-Leware.

Freiberg, J. & Skolouda, S. (2021). Entwurf einer Corporate Sustainability Reporting Directive: Die Zukunft der nichtfinanziellen Berichterstattung. StuB – Unternehmenssteuern und Bilanzen(12), 469–475.

Garcia-Sanchez, I. M., Hussain, N., Khan, S. A., & Martinez-Ferrero, J. (2022). Assurance of corporate social responsibility reports: Examining the role of internal and external corporate governance mechanisms. Corporate Social Responsibility and Environmental Management, 29(1), 89-106. https://doi.org/10.1002/csr.2186.

Gläser-Zikuda, M., Stephan, M., & Hofmann, F. (2022). Qualitative Auswertungsverfahren. In: Reinders, H., Bergs-Winkels, D., Prochnow, A., Post, I. (eds) Empirische Bildungsforschung. Springer VS, Wiesbaden. https://doi.org/10.1007/978-3-658-27277-7_15.

Hahn, R., & Lülfs, R. (2013). Legitimizing Negative Aspects in GRI-Oriented Sustainability Reporting: A Qualitative Analysis of Corporate Disclosure Strategies. Journal Of Business Ethics, 123(3), 401–420. https://doi.org/10.1007/s10551-013-1801-4.

Helfferich, C. (2022). Leitfaden- und Experteninterviews. In: Baur, N., Blasius, J. (eds) Handbuch Methoden der empirischen Sozialforschung. Springer VS, Wiesbaden. https://doi.org/10.1007/978-3-658-37985-8_55.

IAASB (2024). Sustainability Assurance. https://www.iaasb.org/consultations-projects/sustainability-assurance

IAASB (2023). Explanatory Memorandum for Proposed International Standard on Sustainability Assurance (ISSA) 5000. General Requirements for Sustainability Assurance Engagements.

IAASB (2021). Non-Authoritative Guidance on Applying ISAE 3000 (Revised) to Sustainability and Other Extended External Reporting (EER) Assurance Engagements

IFAC, AICPA, & CIMA (2024): The State of Play: Sustainability Disclosure and Assurance. 2019-2022 Trends & Analysis. New York.

IDW Hrsg. (2024): Sustainability Auditor IDW https://www.idw.de/idw-akademie/programmuebersicht/sustainability-auditor-idw/.

IFRS Foundation (2023). Basis for Conclusions on IFRS S1 General Requirements for Disclosure of Sustainabilityrelated Financial Information. https://www.ifrs.org/content/dam/ifrs/publications/amendments/english/2023/issb-2023-c-basis-for-conclusions-on-ifrs-s1-general-requirements-for-disclosure-of-sustainability-related-financial-information-part-c.pdf?bypass=on.

ISAE 3000 (Revised) (2013): Assurance Engagements Other than Audits or Reviews of Historical Financial Information. International Framework for Assurance Engagements and Related Conforming Amendments von Dezember 2013. IAASB, Inter-national Standard on Assurance Engagements.

Jensen, M. C., & Meckling, W. H. (1976): Theory of the Firm: Managerial Behavior, Agency Costs, and Ownership Structure. Journal of Financial Economics, Vol. 3, 305–360. https://doi.org/10.1016/0304-405X(76)90026-X.

Jørgensen, S., Mjøs, A., & Pedersen, L. J. T. (2022). Sustainability reporting and approaches to materiality: tensions and potential resolutions. Sustainability Accounting Management And Policy Journal, 13(2), 341–361. https://doi.org/10.1108/sampj-01-2021-0009.

Knechel, W. R. (2021). The Future of Assurance in Capital Markets: Reclaiming the Economic Imperative of the Auditing Profession. Accounting Horizons, 35(1), 133–151. https://doi.org/10.2308/horizons-19-182

Krueger, P., Sautner, Z., Tang, D. Y., & Zhong, R. (2021). The effects of mandatory ESG disclosure around the world. Journal of Accounting Research. https://doi.org/10.1111/1475-679X.12548.

Kuckartz, U., & Rädiker, S. (2024). Fokussierte Interviewanalyse mit MAXQDA. Schritt für Schritt, 2. Aufl. Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-658-40212-9.

Kühl, S., Stodtholzt, P., & Taffertshofer, A. (2009). Handbuch Methoden der Organisationsforschung, Quantitative und Qualitative Methoden, Wiesbaden, VS, Verl. für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-91570-8.

Lanfermann, G., & Glöckner, A. (2020). Green and More: Sicherung der Zuverlässigkeit nicht-finanzieller Informationen durch eine EU-weite Prüfungspflicht. WPg – Die Wirtschaftsprüfung, 20, 1227–1229.

Lanfermann, G., & Hebestreit, G. (2021). Nachhaltigkeitsberichterstattung – Der neue Richtlinienentwurf der EU-Kommission verschiebt die Kräfteverhältnisse zwischen finanzieller und nicht-finanzieller Berichterstattung: Interview mit Georg Lanfermann, Präsident des DRSC e. V. IRZ – Zeitschrift für Internationale Rechnungslegung, 7, 305–307.

Löffler, J. (2023). Merkmale der modernen Abschlussprüfung. Zugleich ein Plädoyer für den Beruf des Wirtschaftsprüfers. WPg – Die Wirtschaftsprüfung, 7, S. 337–343.

Maas, Karen; Rosendaal, Sanne (2016): Sustainability Targets in Executive Remuneration: Targets, Time Frame, Country and Sector Specification. In: Bus Strat Env 25 (6), S. 390–401. DOI: 10.1002/bse.1880.

Manne, H. (1966). Insider trading and the stock market: New York: The Free Press.

Manne, H. (2005). Insider Trading: Hayek, Virtual Markets, and the Dog that did not Bark. Journal of Corporation Law, Vol. 31: 167–185. http://dx.doi.org/10.2139/ssrn.6796622.

Mayring, P. (2016). Einführung in die qualitative Sozialforschung. Eine Anleitung zu qualitativem Denken. 6. überarbeitete Auflage. Weinheim, Basel: Beltz.

Mayring, Ph., & Fenzl, T. (2022). Qualitative Inhaltsanalyse. In N. Baur & J. Blasius (Hrsg.), Handbuch Methoden der empirischen Sozialforschung. 3. Auflage. Band 1 (691–706). Wiesbaden: Springer VS.

Meyer, Y., & Klose, J. (2023). Green and more: Wesentlichkeitsanalyse als Ausgangspunkt des künftigen Nachhaltigkeitsberichts. WPg – Die Wirtschaftsprüfung, 22, 1225–1284.

Misoch, Sabina. Qualitative Interviews, Berlin, Boston: De Gruyter Oldenbourg, 2019. https://doi.org/10.1515/9783110545982.

Niederberger, M. (2015). Methoden der Experteneinbindung. In: Niederberger, M., Wassermann, S. (eds) Methoden der Experten- und Stakeholdereinbindung in der sozialwissenschaftlichen Forschung. Springer VS, Wiesbaden. https://doi.org/ 10.1007/978-3-658-01687-6_3.

Nishitani, K., Haider, M. B., & Kokubu, K. (2020). Are third-party assurances preferable to third-party comments for promoting financial accountability in environmental reporting?. Journal of cleaner production, 248, 119199. https://doi.org/10.1016/ j.jclepro.2019.119199.

Phillips, H. (2024). Prüfung von Nachhaltigkeitsinformationen – Status quo und Implikationen für die Umsetzung von Vorgaben der CSRD in Deutschland. Der Betrieb, 23–24, S. 1425–1431.

Quick, R. (2024). Verursacht die Prüfungspflicht für Nachhaltigkeitsberichte neue Erwartungslücken? WPg – Die Wirtschaftsprüfung, 6, 301–352.

Quick, R. (2024b). Zur Ausgestaltung der Pflicht zur Prüfung von Nachhaltigkeitsberichten. Anmerkungen zu den im Referentenentwurf des CSRD-UmsG vorgesehenen Regelungen zur Pflichtprüfung. WPg – Die Wirtschaftsprüfung, 9, 461–516.

Quick, R., Toledano, D., & Toledano, J. (2022). Ist eine Prüfungspflicht für Nachhaltigkeitsberichte sinnvoll? Eine wissenschaftsbasierte Diskussion des EU-Vorschlags zu einer überarbeiteten CSR-Richtlinie. IRZ – Zeitschrift für Internationale Rechnungslegung, 1, 3–42.

Reimsbach, D., Hahn, R. & Gürtürk, A. (2018). Integrated Reporting and Assurance of Sustainability Information: An Experimental Study on Professional Investors' 106 Information Processing. European Accounting Review, 27(3), 559–581. https:// doi.org/10.1080/09638180.2016.1273787.

Reineke, R. (2023). Besonderheiten und realökonomische Implikationen der Nachhaltigkeitsberichterstattung. WPg – Die Wirtschaftsprüfung, 12, 684–690.

Richter, C. P., Govenko, W., & Keine, K. (2021). Wie nachhaltig berichten Unternehmen im europäischen Vergleich? Zur Umsetzung der europäischen CSR-Richtlinie. WPg – Die Wirtschaftsprüfung, 11, 703–712.

Schaltegger, S., & Wagner, M. (2010). Sustainable entrepreneurship and sustainability innovation: categories and interactions. Business Strategy And The Environment, 20(4), 222–237. https://doi.org/10.1002/bse.682.

Schäfer, N. & Schöneberger, M. (2021). Green and more: Die EU-Taxonomie – (weit) mehr als drei Kennzahlen. WPg – Die Wirtschaftsprüfung, 10, 633–635.

Schütte, J., & Dickel, J. (2024). Freiwillige Nachhaltigkeitsberichterstattung von KMU. Der geplante VSME-Standard für nicht berichtspflichtige kleine und mittel-große Unternehmen im Kontext von ESRS und DNK. Die Wirtschaftsprüfung, 17, 837–900.

Seele, P., & Gatti, L. (2017): Greenwashing Revisited: In Search for a Typology and Accusation-Based Definition Incorporating Legitimacy Strategies. Business Strategy and the Environment, 26(2), 239–252. https://doi.org/10.1002/bse.1912.

Sopp, K., Baumüller, J., & Scheid, O. (2022). Nachhaltigkeitsberichterstattung.- Nichtfinanzielle Berichterstattung nach dem CSR-RUG und Neuerungen durch die CSRD, 2. Aufl., Herne: NWB.

Staubus, G. J. (2003). An accountant's education. The Accounting Historians Journal, 30(1), 155–196. https://doi.org/10.2308/0148-4184.30.1.155.

Sterzel, J. (2011). Bewertungs- und Entscheidungsrelevanz der Humankapitalberichterstattung – Eine experimentelle Analyse aus der Perspektive privater Anleger. Wiesbaden: Gabler.

Säuberlich, K., & Jordan, S. (2024). Unwesentliche Tochterunternehmen? – Ein Perspektivwechsel im Zuge der Nachhaltigkeitsberichterstattung. WPg – Die Wirtschaftsprüfung, 3, S. 140–146.

Torelli, R., Balluchi, F., & Furlotti, K. (2020). The materiality assessment and stakeholder engagement: A content analysis of sustainability reports. Corporate Social Responsibility And Environmental Management, 27(2), 470–484. https://doi.org/ 10.1002/csr.1813.

United Nations. (2015). Adoption of the Paris Agreement. https://unfccc.int/docu ments/9064

United Nations Global Compact (2021). UN-Global-Compact-Strategy-2021-2023. https://ungc-communications-assets.s3.amazonaws.com/docs/about_the_gc/UN GLOBAL-COMPACT-STRATEGY-2021-2023.pdf.

Velte/Hellriegel (2024). Der Prüfungsvermerk über den Nachhaltigkeitsbericht: ein neues gesetzliches Informationsinstrument. WPg – Die Wirtschaftsprüfung, 18,S. 901–968.

Velte, P. (2023). Auswahl des Prüfungsdienstleisters für Nachhaltigkeitsberichte nach der CSRD. WPg – Die Wirtschaftsprüfung, 23, S. 1285–1348.

Velte, P. (2022).Prüfung der Nachhaltigkeitsberichterstattung durch den Abschlussprüfer Anmerkungen zur geplanten Überarbeitung der CSR-Richtlinie und zum aktuellen Durand Report. WPg – Die Wirtschaftsprüfung, 2, S. 61–112.

Velte, P., & Stawinoga, M. (2022). Prüfung von Nachhaltigkeitsberichten nach der Corporate Sustainability Reporting Directive (CSRD)–Zum Nachhaltigkeitsbericht nach den geplanten ESRS aus der Sicht von Aufsichtsrat und Prüfungsausschuss. WPg: Die Wirtschaftsprüfung (19).

Velte, P. (2021). Prüfung von Nachhaltigkeitsberichten nach dem Entwurf einer „CSR-Richtlinie 2.0“: Vergleichende Analyse der Reformmaßnahmen und kritische Würdigung. WPg – Die Wirtschaftsprüfung, 10, 613–620.

Velte, P. (2021b), Effect of overlapping audit and compensation committee memberships on the readability of management compensation reports in the German HDAX, in: Journal of managerial issues, 33. Jg., S. 27–48.

Velte, P. (2020). Die Lieferkette im Fokus der nichtfinanziellen Berichterstattung: Normative Reichweite, empirische Befunde und Reformdiskussion. DStR – Deutsches Steuerrecht, 37, 1985–2040.

Velte, P., Simon-Heckroth, E., & Borcherding, N. (2020). Zur Notwendigkeit eines „CSR-RUG 2.0“: Eine Bestandsaufnahme empirischer Befunde zur Entscheidungsnützlichkeit der nichtfinanziellen Erklärung. WPg – Die Wirtschaftsprüfung, 22, 1349–1357.

Weber, S. T., & Wernitz, F. (2021) Die Inhaltsanalyse nach Mayring als Auswertungsmethode für wissenschaftliche Interviews. IUBH Discussion Papers – Business & Management, No. 6/2021. Erfurt: IU International University of Applied Sciences.

Weichbold, M. (2022). Pretests. In: Baur, N., Blasius, J. (eds) Handbuch Methoden der empirischen Sozialforschung. Springer VS, Wiesbaden. https://doi.org/10.1007/978-3-658-37985-8_28.

Weuster, C. W., Ottenstein, P., Jost, S., Winter, S., & Zülch, H. (2020). Wie eine externe Prüfung die Transparenz von Nachhaltigkeitsberichten beeinflusst: eine empirische Analyse. Die Wirtschaftsprüfung: WPg, 73(21), 1287–1295.

WPK. (2021). Stellungnahme vom 24. Juni 2021 zum EU-Richtlinienvorschlag zur Nachhaltigkeitsberichterstattung CSRD gegenüber der Europäischen Kommission. https://www.wpk.de/neu-auf-wpkde/csr/.

Yu, E. P-Y., Van Luu, B., & Chen, C. H. (2020). Greenwashing in environmental, social, and governance disclosures. Research in International Business and Finance, 52, 101192. https:// doi. org/ 10. 1016/j. ribaf. 2020. 101192.

Zorio, A., Garcia-Benau, M. A., & Sierra, L. (2013). Sustainability development and the quality of assurance reports: Empirical evidence. Business strategy and the environment, 22(7), 484–500. https://doi.org/10.1002/bse.1764.

3. Vergütungsberichterstattung der HDAX-Unternehmen nach ARUG II: Formale Aspekte und textanalytische Erkenntnisse

3.1 Publikationsdetails

Autoren: Fabian Freches, Jonas Maiwald

Zusammenfassung: Die Einführung des ARUG II hat die Vergütungsberichterstattung börsennotierter Unternehmen umfassend reformiert. Diese Studie analysiert die Auswirkungen auf Umfang und Lesbarkeit der Berichte anhand eines Prä-Post-Vergleichs von 438 Vergütungsberichten der Berichtsjahre 2020 bis 2022. Textanalysen zeigen, dass sich die Lesbarkeit trotz des stark gestiegenen Berichtsumfangs nur geringfügig verschlechtert hat, die Berichte aber weiterhin schwer verständlich bleiben. Viele Unternehmen lassen ihre Berichte freiwillig über die gesetzlichen Mindestanforderungen hinaus prüfen, um ein bewusstes Transparenzsignal zu senden.

Stichwörter: Vergütungsberichte, Textanalyse, ARUG II, Corporate Governance, Post-Implementierungsbewertung

Publikationsstatus: Veröffentlicht in: *Betriebswirtschaftliche Forschung und Praxis (BFuP)* Heft 6 (2025): 585–610

Hinweis: Aufgrund von Formatierungsanpassungen in dieser Dissertation kann es zu geringfügigen Abweichungen gegenüber der veröffentlichten Version kommen.

3.2 Einleitung

Angesichts der Auswirkungen vergangener Finanzkrisen, unzureichender Unternehmenskontrollen sowie einer verstärkten Ausrichtung auf kurzfristige statt langfristige Unternehmensentwicklungen hat die Europäische Union die Aktionärsrichtlinie (Richtlinie (EU) 2017/828) verabschiedet. Diese hebt den Vergütungsbericht als zentrales Instrument hervor, um die Rechenschaftspflicht der Unternehmensleitung auszudehnen. Ziel ist es, Aktionären eine verbesserte Überwachung und Bewertung der Unternehmensführung auf Basis klar definierter Kriterien zu ermöglichen. In Deutschland wurden die Vorgaben der europäischen Aktionärsrichtlinie durch das Gesetz zur Umsetzung der zweiten Aktionärsrechterichtlinie (ARUG II)[1] vom 12.12.2019 maßgeblich mit Formulierung des § 87a AktG und § 162 AktG in nationales Recht überführt.[2]

Die Berichterstattung der Unternehmensorgane wurde somit grundlegend reformiert, denn erstmalig müssen Unternehmen einen nach den in § 162 Abs. 1 Satz 1 AktG definierten Kriterien *klaren* und *verständlichen* Bericht erstellen. Der vorliegende Beitrag untersucht anhand eines Prä-Post-Vergleichs, inwiefern sich die Vergütungsberichterstattung börsennotierter Unternehmen mit der Einführung des ARUG II verändert hat. Neben formalen Aspekten der Berichterstattung richtet die Analyse den Fokus auf die Lesbarkeit der Vergütungsberichte von HDAX-Unternehmen. Die Sprachkomplexität wird mittels Textanalyse systematisch erfasst, um zu bestimmen, inwieweit die gesetzlichen Anforderungen hinsichtlich Klarheit und Verständlichkeit gemäß § 162 Abs. 1 Satz 1 AktG erfüllt werden.

Der Beitrag ist wie folgt strukturiert: Kapitel 3.3 stellt die inhaltlichen Neuerungen des ARUG II zur Vergütungsberichterstattung dar und gibt einen Überblick über empirische Studien zur Vorstandsvergütung sowie textanalytische Ansätze in der Rechnungslegungsforschung. Kapitel 3.4 erläutert die Methodik. Kapitel 3.5 präsentiert die Ergebnisse der Untersuchung, ordnet sie ein und interpretiert sie unter Berücksichtigung des gewählten methodischen Vorgehens. Kapitel 3.7 fasst die zentralen Erkenntnisse zusammen.

3.3 Hintergrund

3.3.1 Gesetz zur Umsetzung der zweiten Aktionärsrechterichtlinie (ARUG II)

Die gesetzlichen Vorgaben zum Vergütungsbericht sind in § 162 AktG normiert. Demnach ist ein separater Vergütungsbericht für ab dem 1.1.2021 beginnende Geschäftsjahre von Vorstand sowie

[1] Vgl. *Bundesministerium der Justiz und für Verbraucherschutz* (2019).
[2] Vgl. *Kolb et al.* (2022), S. 33-35.

Aufsichtsrat jeder börsennotierten Gesellschaft jährlich zu erstellen.[3] Im Gegensatz zu den vorherigen handelsrechtlichen Vorgaben ist die Erstellung des Vergütungsberichts als eine Gemeinschaftsaufgabe beider Unternehmensorgane zu verstehen.[4] In den sachlichen Anwenderkreis fallen sämtliche börsennotierten Gesellschaften in der Rechtsform einer Aktiengesellschaft (*AG*), Europäische Aktiengesellschaften (*SE*) mit Sitz in Deutschland sowie Kommanditgesellschaften auf Aktien (*KGaA*)[5], wobei Letztere nur mit inhaltlichen Einschränkungen einbezogen werden.[6] Eine verpflichtende Anwendung des § 162 AktG ist über eine entsprechende Regelung in der Unternehmenssatzung auch für nicht börsennotierte Unternehmen denkbar.[7]

Im Vergütungsbericht sind für den relevanten Personenkreis des Vorstands und des Aufsichtsrats nach den Vorgaben des § 162 Abs. 1 Satz 2 AktG unter anderem Angaben zu allen festen und variablen Vergütungsbestandteilen (Nr. 1), zur vergleichenden Darstellung der jährlichen Veränderungen der Vergütung (Nr. 2), zu den gewährten und zugesagten Aktien sowie Aktienoptionen (Nr. 3) und etwaigen Abweichungen vom Vergütungssystem (Nr. 5) zu machen. Ferner sind etwa Angaben nach § 162 Abs. 2 Satz 1 AktG auszuweisen, sofern einem Vorstandsmitglied von einem Dritten im Hinblick auf seine Tätigkeit als Vorstandsmitglied Leistungen zugesagt oder im Geschäftsjahr gewährt worden sind (Nr. 1) oder für den Fall der vorzeitigen Beendigung seiner Tätigkeit zugesagt worden sind (Nr. 2). Eine weitere zentrale Neuerung ist der Vertikalvergleich, bei dem die Organvergütung mit der Ertragsentwicklung der Gesellschaft und der Arbeitnehmervergütung in Relation gesetzt wird.[8] Grundsätzlich dürfen Angaben jedoch unterlassen werden, sofern sie nach vernünftiger kaufmännischer Beurteilung der Gesellschaft einen nicht unerheblichen Nachteil zufügen können.[9]

Die handelsrechtlichen Berichtspflichten, insbesondere die Anhangangaben gemäß § 285 Nr. 9 HGB, bleiben weiterhin bestehen, wurden jedoch speziell für börsennotierte Unternehmen im Hinblick auf die Vermeidung von Redundanzen angepasst.[10] Ferner sind die Neuerungen des ARUG II hinsichtlich der Berichterstattung in der gegenwärtigen Fassung des Deutschen Corporate Governance Kodex (DCGK) als Grundsatz mitaufgenommen worden.[11] Ansonsten bestehen keine freiwillig

[3] Vgl. § 162 Abs. 1 AktG; § 26j Abs. 2 AktGEG; *MüKoAktG/Spindler*, AktG § 162 Rn. 91; *Rimmelspacher/Roland* (2020), S. 202 f.
[4] Vgl. *MüKoAktG/Spindler*, AktG § 162 Rn. 12.
[5] Mangels Vorstands sind im Falle einer börsennotierten Kommanditgesellschaft auf Aktien der Komplementär sowie der Aufsichtsrat zur Erstellung des Vergütungsberichts verpflichtet. Vgl. *MüKoAktG/Spindler*, AktG § 162 Rn. 12.
[6] Vgl. *MüKoAktG/Spindler*, AktG § 162 Rn. 11; *Rimmelspacher/Roland* (2020), S. 201 f.
[7] Vgl. *Grigoleit/Rachlitz*, AktG § 162 Rn. 27; *Kuthe/Reuters* (2022), S. 2115.
[8] Vgl. *Rimmelspacher/Kliem* (2021), S. 1000.
[9] Vgl. § 162 Abs. 6 AktG.
[10] Vgl. *Beck Bil-Komm/Grottel*, HGB § 285 Rn. 260.
[11] Vgl. *Regierungskomission Deutscher Corporate Governance Kodex* (2022), Grundsatz 26.

anzuwendenden Empfehlungen und Anregungen im Kodex. Demgemäß sind die Vorgaben aus dem Aktiengesetz alleinige Grundlage zur Erstellung des Vergütungsberichts.[12]

Der Vergütungsbericht soll *klar* und *verständlich* über die im letzten Geschäftsjahr jedem einzelnen gegenwärtigen oder früheren Mitglied des Vorstands und des Aufsichtsrats der Gesellschaft und von Unternehmen desselben Konzerns gewährte und geschuldete Vergütung informieren.[13] Die Formulierung durch den deutschen Gesetzgeber orientiert sich dabei an dem durch die EU-Kommission veröffentlichten Entwurf einer Leitlinie zur Vergütungsberichterstattung.[14] Demnach muss der Bericht generell klar, prägnant, sinnvoll und verständlich sein.[15] Die normativen Empfehlungen auf EU-Ebene sind somit detaillierter als die nationale Umsetzung. Wenngleich die Leitlinien keinen rechtsverbindlichen Charakter haben, so sind die Empfehlungen bei Auslegungsfragen zu berücksichtigen.[16] Die Darstellungsvorgabe der Klarheit und Verständlichkeit gilt gleichermaßen für den Vergütungsbericht nach § 162 AktG als auch für das Vergütungssystem nach § 87a AktG. Aus den unbestimmten Rechtsbegriffen der Klarheit und Verständlichkeit ergibt sich für die Unternehmensorgane somit ein Spielraum für die Gestaltung des Vergütungsberichts.[17] Vorstand und Aufsichtsrat können sich an dem dem Vergütungsbericht zugrunde liegenden Empfängerhorizont des Adressatenkreises orientieren.[18] So wird in der Regierungsbegründung weiter formuliert, dass der Vergütungsbericht „[…] *dem durchschnittlich informierten, situationsadäquat aufmerksamen und verständigen Aktionär eine informierte Entscheidung* […]“ ermöglichen soll.[19] Mitunter wird der Empfängerhorizont auch mit einem allgemeinen, mit den Verhältnissen von Aktiengesellschaften in den Grundzügen vertrauten Leser definiert, während klargestellt wird, dass nicht jede einzelne Person zum Adressatenkreis zählt.[20]

Grundsätzlich sind nach Ansicht des deutschen Gesetzgebers langatmige Formulierungen mit technischen Begriffen zu unterlassen, während hingegen Beispiele und Schaubilder den Zweck der Klarheit und Verständlichkeit fördern können.[21] Im Gegensatz dazu weist die EU-Kommission jedoch darauf hin, dass erläuternde schriftliche Ausführungen insbesondere für komplexe Sachverhalte

[12] Vgl. *Rimmelspacher/Roland* (2020), S. 201 f.
[13] Vgl. § 162 Abs. 1 AktG.
[14] Vgl. *EU-Kommission* (2022).
[15] Vgl. *EU-Kommission* (2022), S. 2.
[16] Vgl. *MüKoAktG/Spindler*, AktG § 162 Rn. 31.
[17] Vgl. *Hölters/Weber*, AktG § 162 Rn. 13; *MüKoAktG/Spindler*, AktG § 162 Rn. 32.
[18] Vgl. *Anzinger* (2019), S. 84.
[19] Vgl. *Deutscher Bundestag* (2019), S. 109.
[20] Vgl. *Grigoleit/Rachlitz*, AktG § 162 Rn. 17.
[21] Vgl. *Deutscher Bundestag* (2019), S. 109.

dienlich sein können.[22] Ferner muss der Vergütungsbericht alle Pflichtinformationen enthalten, wenngleich Negativmeldungen nicht angezeigt werden müssen. Verweise auf andere Veröffentlichungen sind dabei für Pflichtangaben unzulässig, während der Gesetzgeber den Verweis auf über die Berichtspflicht hinausgehende Informationen gestattet.

Die Berichtsgrundlage für den Vergütungsbericht wird durch die ebenfalls nach ARUG II reformierten Vorgaben für das Vergütungssystem von Vorständen börsennotierter Gesellschaften determiniert. Die Klarstellung einer auf die nachhaltige *und* langfristige Entwicklung der Gesellschaft ausgerichteten Vergütungsstruktur stellt eine zentrale Neuerung dar. Sie basiert auf den Grundsätzen für die Bezüge von Vorstandsmitgliedern nach § 87 AktG sowie den neu eingeführten Vorgaben für die Vergütungssysteme börsennotierter Gesellschaften nach § 87a AktG, die als Operationalisierung dieser Grundsätze dienen. Hierdurch wird klargestellt, dass der Nachhaltigkeitsaspekt nicht nur auf eine zeitliche Dimension im Sinne der alten Fassung auszulegen ist, sondern auch ausdrücklich nichtfinanzielle Zielvorgaben berücksichtigt werden müssen, womit der Gesetzgeber die gesellschaftliche Verantwortung der Unternehmen betont.[23] In der Praxis etabliert sich eine klare Tendenz zur Einbindung von ESG-Faktoren in die Vorstandsvergütung.[24] Zusätzlich sind im neu eingeführten § 87a AktG die Mindestinhalte zur Vergütung von Vorstandsmitgliedern geregelt. So muss nach § 87a Abs. 1 AktG das Vergütungssystem Angaben u. a. zu der Maximalvergütung (Nr. 1), zu dem Beitrag der Vergütung zur Förderung der Geschäftsstrategie und zur langfristigen Entwicklung der Gesellschaft (Nr. 2), zu allen festen und variablen Vergütungsbestandteilen (Nr. 3) sowie zu allen finanziellen und nichtfinanziellen Leistungskriterien für die Gewährung variabler Vergütungsbestandteile (Nr. 4) beinhalten.[25] Mit Einführung des ARUG II stimmt die Hauptversammlung gemäß § 120a Abs. 1 AktG mindestens alle vier Jahre über das Vergütungssystem ab und muss den Vergütungsbericht nach § 120a Abs. 4 AktG jährlich billigen. Wenngleich das Votum der Hauptversammlung allenfalls beratenden Charakter hat, so ist davon auszugehen, dass das Meinungsbild der Aktionäre notwendigerweise durch die Sorgfaltspflichten des Aufsichtsrats nach § 116 AktG zu berücksichtigen sein wird.[26] Ferner können die sogenannten Say-on-pay-Mechanismen für mediale Aufmerksamkeit sorgen und zum Handeln zwingen.[27]

[22] Vgl. *EU-Kommission* (2022), S. 4.
[23] Vgl. *Velte* (2020b), S. 14.
[24] Vgl. etwa *Tönnies/Tönnies* (2024); *Schmidt/Pacher* (2022); für Unternehmen im EURO STOXX 50 *Mühlberger/Woern* (2024).
[25] Vgl. *Stöber* (2020), S. 94.
[26] Vgl. *MüKoAktG/Spindler*, AktG § 87a Rn. 8; *MüKoAktG/Kubis*, AktG § 120a Rn. 22.
[27] Vgl. etwa *Manager Magazin* (2022). Für empirische Befunde zur Wirkung der Say-on-pay-Mechanismen siehe *Balsam et al.* (2016).

Der Vergütungsbericht ist von einem Wirtschaftsprüfer dahingehend formell zu prüfen, ob die vom Gesetz geforderten Inhalte auch tatsächlich veröffentlicht wurden.[28] Der Wirtschaftsprüfer testiert die Berichterstattung dann in einem separaten Vermerk, der dem Vergütungsbericht beizufügen ist.[29] Eine inhaltliche Prüfung bleibt hingegen freiwillig, wenngleich sie vom Berufsstand gefordert wird, um der Gefahr einer Vergrößerung der Erwartungslücke entgegenzutreten.[30] Die Prüfung des Vergütungsberichts stellt somit einen eigenständigen Prüfungsgegenstand dar. Unbeachtlich dessen hat der Wirtschaftsprüfer im Rahmen der Jahresabschlussprüfung nach den handelsrechtlichen Vorgaben zu prüfen, ob die berichtspflichtige Gesellschaft in der Erklärung zur Unternehmensführung den geforderten Verweis nach § 289f Abs. 2 HGB auf den Vergütungsbericht aufgenommen hat.[31] Nach Beschluss der Hauptversammlung sind der Vergütungsbericht sowie der Prüfungsvermerk unverzüglich auf der Internetseite der Gesellschaft zu veröffentlichen und für die Dauer von zehn Jahren kostenfrei zugänglich zu machen.[32]

3.3.2 Deskriptive Studien zur Vorstandsvergütung im deutschsprachigen Raum

Regulatorische Veränderungen an den Vorgaben zur Organvergütung sind regelmäßig Gegenstand empirischer Untersuchungen.[33] Eine Untersuchung zur Entwicklung der Vorstandsvergütung in Deutschland im Kontext des Gesetzes zur Angemessenheit der Vorstandsvergütung (VorstAG), einer Vorgängerregelung des ARUG II, zeigt, dass die durchschnittlichen Vorstandsgehälter im Zeitraum von 2006 bis 2018 lediglich moderat von 1,83 Millionen € auf 1,99 Millionen € anstiegen. Zudem gewinnen langfristige Vergütungsbestandteile, etwa Aktienoptionen, zunehmend an Bedeutung in der Gesamtvergütung börsennotierter Unternehmen in Deutschland und ältere CEOs verdienen im Mittel mehr als jüngere CEOs.[34] Ferner stellen *Tröger/Walz* fest, dass die mit dem VorstAG erstmalig eingeführten Say-on-pay-Mechanismen zur Abstimmung über die Vergütungssysteme nur im begrenzten Maße eine Wirkung entfalten haben.[35] *Aust* bewertet mithilfe eines Scoring-Modells die Qualität der Vergütungsberichte börsennotierter Unternehmen für das Geschäftsjahr 2016 bzw. 2016/2017. Sie stellt fest, dass die Vergütungsberichterstattung der DAX-30-Unternehmen in den

[28] Das Institut der Wirtschaftsprüfer (IDW) hat mit der Verabschiedung eines entsprechenden Prüfungsstandards (IDW PS 870) auf die formelle Pflichtprüfung reagiert.
[29] Vgl. § 162 Abs. 3 AktG.
[30] Vgl. *IDW* (2018), S. 4.
[31] Vgl. *Kolb et al.* (2022), S. 188 f.
[32] Vgl. § 162 Abs. 4 AktG; *MüKoAktG/Spindler*, AktG § 162 Rn. 92.
[33] Vgl. etwa *Götz/Friese* (2010) für die Wirkung des Vorstandsvergütungsangemessenheitsgesetzes oder *Burg et al.* (2016) zur Wirkung von HGB, DCGK und der IFRS auf die Vergütungspraxis.
[34] Vgl. *Beck et al.* (2020), S. 790, 802, 806.
[35] Vgl. *Tröger/Walz* (2019), S. 381.

Bereichen Transparenz, Nachvollziehbarkeit und Kommunikation qualitativ höherwertiger ist als die der MDAX-Unternehmen.[36]

Bisherige Studien, die den Effekt der ARUG-II-Regulierung untersuchen, konzentrieren sich auf die inhaltliche Gestaltung der Vergütungssysteme. *Tönnies/Tönnies* nutzen ein Scoring-Modell und untersuchen die Relevanz von ESG-Faktoren in den Vergütungssystemen. Dabei verwendet die überwiegende Mehrheit der DAX-40- sowie SDAX-Unternehmen ESG-Faktoren in der Vorstandsvergütung. In der Qualität der Berichterstattung zeigen sich deutliche Index- und Branchenunterschiede, was auf ein heterogenes Berichtsverhalten der Unternehmen hindeutet.[37] Allerdings stellen *Bannier/Reinschmidt* fest, dass Nachhaltigkeitsziele eher in der kurzfristigen variablen Vergütung berücksichtigt werden. Dies lässt die Hypothese zu, dass Nachhaltigkeitsziele zwar im Sinne gesetzlicher Vorgaben berücksichtigt werden, jedoch geringere Bedeutung erhalten und nicht als wesentlicher Anreiz für eine langfristige positive Unternehmensentwicklung betrachtet werden. Andernfalls wäre eine stärkere Gewichtung in der langfristigen variablen Vergütung zu erwarten.[38] Hinsichtlich der gewährten Vorstandsvergütung pro Mitglied in Gänze lässt sich im Median kein eindeutiger Trend erkennen, wenngleich dies unter Berücksichtigung der langjährigen variablen Vergütung noch zu bestätigen ist.[39] Der Umfang der Vergütungsberichte von DAX-Unternehmen ist von durchschnittlich 19 auf 31 Seiten gestiegen, was als Ausdruck eines zunehmenden Komplexitätsbedarfs in der Berichterstattung interpretiert wird.[40] Weiterführende Studien hinsichtlich der Eigenschaften des Vergütungsberichts als eigenständiges Berichtsinstrument fehlen im deutschsprachigen Raum und begründen somit den weiteren Forschungsbedarf.

3.3.3 Textanalytische Studien in der Rechnungslegungsforschung und theoretische Fundierung

Die Publizitätspflicht von Unternehmen ermöglicht grundsätzlich die Analyse der offenzulegenden Dokumente durch Textanalysen. Als Textanalyse wird im erweiterten Sinne die Anwendung einer Natural Language Processing-Methode (NLP) auf Textdaten zur automatischen Informationsextraktion oder -messung verstanden.[41] Die Textanalyse findet vermehrt Anwendung in

[36] Vgl. *Aust* (2019), S. 468.
[37] Vgl. *Tönnies/Tönnies* (2024), S. 232.
[38] Vgl. *Bannier/Reinschmidt* (2022), S. 640.
[39] Vgl. *Armeli/Misterek* (2023), S. 1; *Götz/Stahl* (2023), S. 196.
[40] Vgl. *Götz et al.* (2022), S. 219.
[41] Vgl. *Bochkay et al.* (2023), S. 766. Für einen ausführlichen Literaturüberblick von NLP-Methoden in der rechnungslegungsbezogenen Forschung siehe *Loughran/McDonald* (2016) und *Bochkay et al.* (2023). *Kang et al.* (2020) geben einen umfassenden Literaturüberblick für die Managementforschung.

der Rechnungslegungsforschung, um die Ähnlichkeit[42], die Stimmung[43] oder die Lesbarkeit[44] von Textdokumenten zu erheben.

Die Lesbarkeitsanalyse hat sich in der Rechnungslegungsforschung zur Untersuchung der Verständlichkeit sowie zur Identifikation potenzieller Manipulationen durch bewusste sprachliche Steuerung von Wahrnehmung und Vertrauen der Berichtsadressaten etabliert. Die Forschungsansätze unterscheiden sich dabei sowohl in der Zielsetzung – deskriptiv versus kausal – als auch in der Wahl der Berichtsinhalte (z. B. Management Discussion and Analysis (MD&A), CSR-Berichte, Risiko- und Chancenberichte).[45] Empirische Studien belegen, dass Unternehmensleitungen gezielt komplexe Berichtsdokumente einsetzen können, um suboptimale Unternehmensentwicklungen zu kaschieren.[46] So zeigt etwa eine breite empirische Untersuchung, dass Unternehmen mit schwacher Performance tendenziell weniger lesbare Jahresberichte (Form 10-K) veröffentlichen, während eine höhere Lesbarkeit im MD&A-Abschnitt signifikant mit größerer Gewinnpersistenz einhergeht – ein Befund, der auf potenziell manipulierende Sprachgestaltung (obfuscation) hindeutet.[47] So kann die Einschätzung von Investoren durch die Lesbarkeit der Unternehmenspublikationen gelenkt werden.[48]

Methodisch dominieren in der Literatur quantitative Lesbarkeitsindizes wie der Flesch Reading Ease (FRE), Gunning Fog Index (FOG), Lesbarkeitsindex nach Björnsson (LIX) oder die Wiener Sachtextformel (WSTF), mitunter angepasst an deutschsprachige Texte (z. B. Flesch Reading Ease nach Amstad (FRE_{Amstad})).[49] Teilweise werden auch pragmatische Lesbarkeitsmetriken der Textkomplexität verwendet, etwa die Anzahl der Wörter[50] oder – bei standardisierten digitalen Berichten – die Dateigröße als Indikatoren für Informationsdichte und Umfang.[51]

Die Lesbarkeitsindizes ermöglichen skalierbare Auswertungen und zeichnen sich durch ihre Objektivierbarkeit und Vergleichbarkeit bei großen Datensätzen aus, erfassen jedoch primär formale Merkmale wie Satzlänge oder Silbenanzahl – nicht jedoch semantische Verständlichkeit, argumentative Struktur oder kontextuelle Logik.[52] Die strukturelle Textanalyse hingegen verfolgt

[42] Vgl. *Mauritz et al.* (2023); *Quick et al.* (2021).
[43] Vgl. *Lohmann/Ohliger* (2020); *Lohmann/Danisch* (2021); *Pöferlein* (2021); *Frankel et al.* (2022).
[44] Vgl. *Guay et al.* (2016); *Bonsall et al*. (2017); *Velte* (2020a); für einen umfassende Literaturübersicht siehe *Gosselin et al*. (2021).
[45] Vgl. *Gosselin et al.* (2021), S. 553-557.
[46] Vgl. *Loughran/McDonald* (2014), S. 1648.
[47] Vgl. *Li* (2008), S. 225-247.
[48] Vgl. *El-Sayed et al.* (2021); *Gao et al.* (2023); *Quick et al*. (2024).
[49] Vgl. *Gosselin et al.* (2021), S. 558 f.
[50] Vgl. *You/Zhang* (2009).
[51] Vgl. *Loughran/McDonald* (2014).
[52] Vgl. *Li* (2008), S. 222; *Gosselin et al.* (2021), S. 558 f.

einen funktional-diagnostischen Ansatz. Sie ermöglicht nicht nur die Bewertung sprachlicher Komplexität, sondern auch die gezielte Identifikation sogenannter Nullpositionen – das heißt inhaltlich erwartbare, aber ausgelassene Informationen. Damit erlaubt die strukturelle Textanalyse Rückschlüsse auf kommunikative Intentionen und potenzielle Transparenzvermeidung.[53]

Lesbarkeitsindizes dienen somit der quantitativen Erfassung sprachlicher Komplexität und damit der kognitiven Belastung beim Textverständnis – beeinflusst etwa durch Satzstruktur und Wortverwendung.[54] Im Folgenden werden die in dieser Untersuchung eingesetzten Messansätze näher erläutert.

Der FOG sowie der FRE eignen sich für die Analyse englischsprachiger Dokumente. Beide werden wie folgt ermittelt:[55]

$$FOG^{56}=0{,}4\left[\left(\frac{\text{Anzahl Wörter}}{\text{Anzahl Sätze}}\right)+100\left(\frac{\text{Anzahl komplexer Wörtern}}{\text{Anzahl Wörter}}\right)\right]$$

Hierbei gelten Wörter mit mehr als zwei Silben als komplex.

$$FRE^{57}=206{,}835-1{,}015\left(\frac{\text{Anzahl Wörter}}{\text{Anzahl Sätze}}\right)-84{,}6\left(\frac{\text{Anzahl Silben}}{\text{Anzahl Wörter}}\right)$$

Ein niedriger FOG weist auf einen leicht lesbaren Text hin, während ein hoher Wert auf eine schwer verständliche Textstruktur hindeutet. Im Gegensatz dazu signalisiert ein niedriger (hoher) Wert des FRE einen komplexen (einfachen) Text.[58] Der FOG misst Verständlichkeit anhand von Satzlänge und Wortkomplexität. Technische Bücher (19,5) sind besonders anspruchsvoll, während Jugendmagazine (8,6) leicht verständlich sind. So entspricht ein Wert von 17 dem Leselevel eines Hochschulabsolventen, während ein Wert von 6 einem Sechstklässler entspricht.[59] Der FRE-Index skaliert die Lesbarkeit von 0 bis 100. Texte unter 30 sind sehr schwer verständlich (akademisches Niveau), während Werte über 90 sehr leichte Texte kennzeichnen, die bereits ab der 4. Schulklasse lesbar sind.[60]

[53] Vgl. *Grottke* (2009).
[54] Vgl *Loughran/McDonald* (2014), S. 1647; *Bochkay et al.* (2023), S. 779.
[55] Vgl. *Loughran/McDonald* (2016), S. 1193; *Stone/Parker* (2013), S. 35.
[56] Vgl. *Gunning* (1952).
[57] Vgl. *Flesch* (1948).
[58] Vgl. *Loughran/McDonald* (2014), S. 1648; *Stone/Parker* (2013), S. 35.
[59] Vgl. *Courtis* (1995), S. 7.
[60] Vgl. *Courtis* (1995), S. 7; *Flesch* (1948).

Der LIX sowie der FRE_{Amstad} eignen sich zur Ermittlung der Lesbarkeit deutschsprachiger Texte.[61] Hier lauten die Formeln zur Erhebung der Lesbarkeit:

$$LIX^{62} = \left(\frac{\text{Anzahl langer Wörter}}{\text{Anzahl Wörter}}\right) 100 + \left(\frac{\text{Anzahl Wörter}}{\text{Anzahl Sätze}}\right)$$

Hierbei gelten Wörter mit mehr als sechs Buchstaben als lang.

$$FRE_{Amstad}{}^{63} = 180 - \frac{\text{Anzahl Wörter}}{\text{Anzahl Sätze}} - \left(58{,}5\,\frac{\text{Anzahl Silben}}{\text{Anzahl Wörter}}\right)$$

Der LIX-Wert misst die Lesbarkeit anhand der Anzahl langer Wörter und Sätze. Texte mit einem Wert über 60 gelten als sehr schwer verständlich, während Werte zwischen 50 und 55 als schwer und zwischen 40 und 45 als durchschnittlich verständlich eingestuft werden. Texte unter 25 sind besonders leicht lesbar.[64] Dementgegen signalisiert ein niedriger (hoher) Wert des FRE_{Amstad} einen komplexen (einfachen) Text. Der FRE_{Amstad}, eine Anpassung des klassischen FRE-Index für deutschsprachige Texte, verwendet eine Skala, bei der Werte unter 20 für sehr schwierige Texte stehen, während Werte zwischen 20 und 30 als schwierig und zwischen 30 und 40 als anspruchsvoll gelten.[65]

Die Analyse der Lesbarkeit in Vergütungsberichten lässt sich durch die Einbeziehung theoretischer Modelle fundiert begründen. Die Impression-Management-Theorie (bzw. Obfuscation Theory) geht davon aus, dass die Unternehmensführung die Berichterstattung gezielt so gestaltet, dass bestimmte Informationen gegenüber den Adressaten verschleiert oder versteckt werden.[66] Das übergeordnete Ziel besteht darin, die Wahrnehmung der Adressaten vom Unternehmen gezielt zu beeinflussen.[67] Eine derart manipulierte Darstellung erfolgt nicht nur durch das Weglassen nachteiliger Informationen, die etwa durch strukturelle Textanalyse identifiziert werden kann, sondern auch durch eine sprachlich komplexe Ausgestaltung, bei der negative Inhalte zwar enthalten sind, jedoch im Gesamtduktus schwer identifizierbar oder interpretierbar bleiben. Vor diesem Hintergrund ermitteln Lesbarkeitsindizes das Komplexitätsniveau und können mögliche Impression-Management-Techniken aufdecken.

[61] Vgl. *Diederich/Velte* (2019); *Heyd/Anselmann* (2019); *Heilmann et al.* (2020); *Gros/Hanke* (2021); *Quick et al.* (2023), *Quick et al.* (2024).
[62] Vgl. *Björnsson* (1968); siehe ferner *Heilmann et al.* (2020, S. 711), die die Anwendung des LIX anhand eines konkreten Beispiels veranschaulichen.
[63] Vgl. *Amstad* (1978), S. 80.
[64] Vgl. *Anderson* (1983), S. 494; *Courtis* (1995), S. 8; *Björnsson* (1983), S. 484.
[65] Vgl. *Amstad* (1978), S. 80 & S. 153.
[66] Vgl. *Brennan et al.* (2009), S. 790; *Gosselin et al.* (2021), S. 563.
[67] Vgl. *Hohmann/Schäfer* (2020), S. 461.

Hinsichtlich der Vergütungsberichterstattung besteht aus Sicht des Managements ein theoretischer Anreiz zur Verschleierung insbesondere dann, wenn die vergütungsrelevanten Ziele zwischen Vorstand und Adressaten auseinanderfallen. Solche Zieldivergenzen lassen sich in unterschiedlichen Konstellationen argumentieren.[68] So verfolgen Vorstand und Aktionäre zwar grundsätzlich das übergeordnete Ziel der Gewinnmaximierung. Allerdings können sich die Interessen im Hinblick auf die Nachhaltigkeit der Gewinne deutlich unterscheiden: Während der Vorstand primär an der Optimierung kurzfristiger variabler Vergütung interessiert sein kann – beispielsweise durch den Ausschluss aufwandswirksamer Sachverhalte mittels Non-GAAP-Kennzahlen oder bilanzpolitischer Maßnahmen – verfolgen Aktionäre in der Regel eine nachhaltige Unternehmenswertsteigerung, die typischerweise über langfristige variable Vergütungskomponenten incentiviert wird. Wird der Adressatenkreis über die Aktionäre hinaus erweitert, so ist ebenfalls denkbar, dass etwa ESG-relevante Vergütungsziele nur symbolisch integriert und nicht ernsthaft verfolgt und somit in Gänze die Interessen der Stakeholder vernachlässigt werden. Die bisherige Forschung belegt, dass die Vergütungsstruktur Vorstände dazu bewegen kann, die Lesbarkeit der Berichterstattung aus persönlich-opportunistischen Gründen zu erschweren.[69]

3.4 Methodik

3.4.1 Untersuchungsgegenstand und Sample

Die vorliegende Studie analysiert die Vergütungsberichte der HDAX-Unternehmen, die zum Stichtag 31.12.2022 über die Geschäftsjahre 2020 bis 2022 hinweg kontinuierlich im Index vertreten waren. Zur Untersuchung der Auswirkungen des ARUG II wurde ein Prä-Post-Vergleich durchgeführt. Dabei repräsentieren die Vergütungsberichte für das Geschäftsjahr 2020 die letzte Berichtsperiode vor Inkrafttreten des ARUG II, während die Vergütungsberichte ab dem Geschäftsjahr 2021 die erste Berichtsperiode nach Inkrafttreten des Gesetzes darstellen.[70] Im Falle abweichender Geschäftsjahre wurde sichergestellt, dass Vergütungsberichte, die gemäß der Ergebnisdarstellung als 2021 klassifiziert wurden, dem ersten Berichtsdurchgang der neuen Gesetzgebung entsprechen.[71]

[68] Solche Zielkonflikte bilden ebenfalls den zentralen Gegenstand der Agency-Theorie, die davon ausgeht, dass aufgrund asymmetrischer Informationsverteilung zwischen Prinzipal und Agent Anreize zu opportunistischem Verhalten entstehen.
[69] Vgl. *Li* (2008); *Laksmana et al.* (2012).
[70] Da die Neuregelungen des ARUG II ausschließlich für börsennotierte Unternehmen gelten, wurden Vergütungsangaben nicht kapitalmarktorientierter Unternehmen von der Untersuchung ausgeschlossen.
[71] In der Stichprobe weisen sieben Unternehmen ein abweichendes Geschäftsjahr auf. So führt bspw. die Aurubis AG in ihrem Vergütungsbericht, eingebettet im Geschäftsbericht 2020/2021 (Geschäftsjahr 1.10.2020 bis 30.9.2021), explizit an, dass dieser nicht nach den Vorgaben des ARUG II erstellt wurde. Folglich wurde der darauffolgende Vergütungsbericht für das Geschäftsjahr 2021/2022 in die Ergebnisse für 2021 aufgenommen, da dieser erstmals gemäß § 162 AktG erstellt wurde. Einige Unternehmen haben wiederum die neuen Vorgaben bereits vorzeitig umgesetzt, obwohl die rechtliche Verpflichtung erst im folgenden Geschäftsjahr greifen würde. Ein Beispiel hierfür sind die Siemens AG und die Siemens

Ausgehend von 100 Unternehmen, die zum Stichtag 31.12.2022 im HDAX geführt wurden, sind 19 Unternehmen, die während des Beobachtungszeitraums nicht kontinuierlich im HDAX gelistet waren, ausgeschlossen worden, um die Konsistenz des Prä-Post-Vergleichs zu wahren. Aus der Grundgesamtheit wurden zusätzlich acht Unternehmen mit ausländischem Hauptsitz ausgeschlossen. Im Sinne guter Investor Relations sollten die Vergütungsberichte neben der deutschen auch in der englischen Sprache veröffentlicht und entsprechend gesammelt werden. Die endgültige Stichprobe umfasst somit 73 Unternehmen und 219 deutsch-/englischsprachige Vergütungsberichte. Die Branchenzuordnung erfolgte anhand der Klassifizierung der Deutschen Börse.[72] *Abbildung 3.1* gibt eine Übersicht der Branchenverteilung in der finalen Stichprobe.

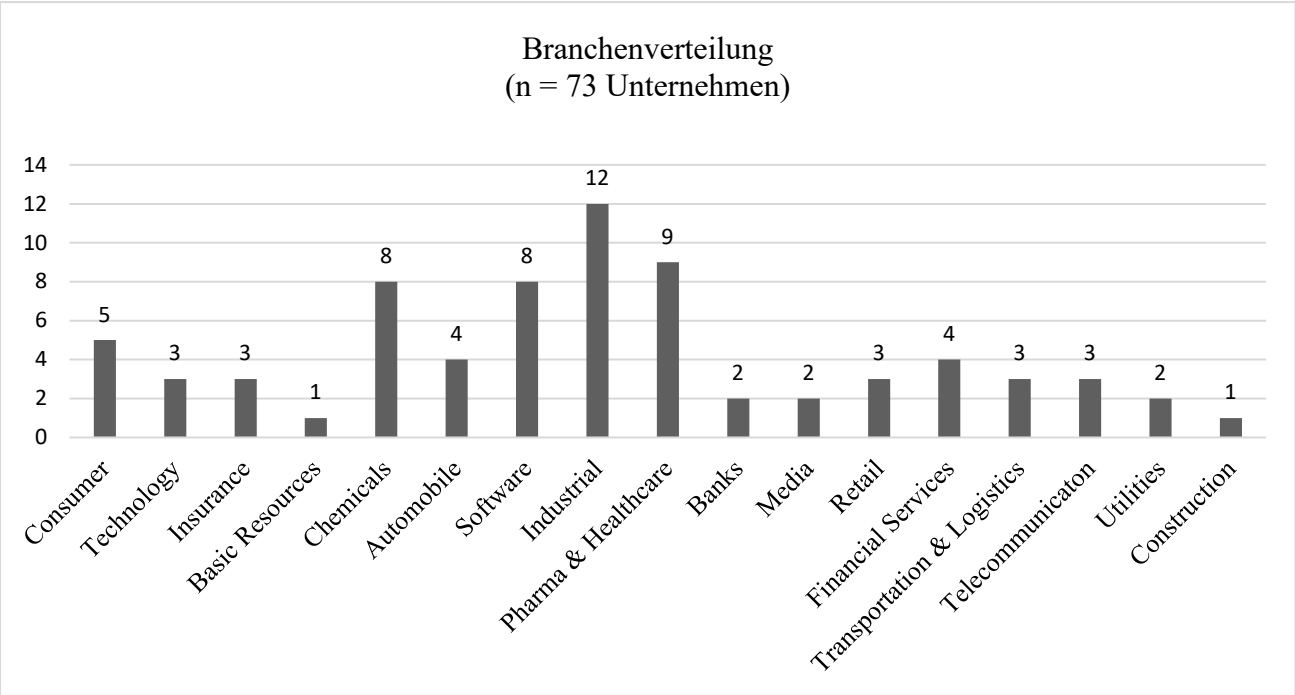


Abbildung 3.1: Branchenverteilung der Unternehmen

Die Unternehmen weisen zum 31.12.2022 in Summe ein Marktkapital von 1.457,66 Milliarden € auf und repräsentieren damit 75,45 % des gesamten Marktkapitals der HDAX-Unternehmen zum Stichtag.

3.4.2 Datenerhebung und Aufbereitung

Die Vergütungsberichte der Geschäftsjahre 2020 bis 2022 wurden händisch erhoben und hinsichtlich ihrer Verfügbarkeit als eigenständige Berichte, als Bestandteil des Geschäftsberichts oder in

Healthineers AG, die ihre Vergütungsberichte 2021 für das Geschäftsjahr vom 1.10.2020 bis zum 30.9.2021 bereits nach den Regelungen des ARUG II erstellt haben und damit über die gesetzlichen Anforderungen hinausgehen.

[72] Vgl. *Deutsche Börse AG* (2024).

alternativen Veröffentlichungsformen geprüft.[73] Notwendige Schritte, wie die Entschlüsselung und Extraktion der Berichte aus PDF-Dokumenten, wurden durchgeführt. Zur Optimierung der maschinellen Verarbeitung erfolgte eine Konvertierung der PDF-Dokumente in Word-Formate, um die anschließende Datenbereinigung effizienter zu gestalten. Die Bereinigung der Berichte erfolgte durch eine Kombination aus maschinellen und manuellen Schritten. Ziel war es, Missinterpretationen durch die angewandten Lesbarkeitsalgorithmen zu minimieren und die orthografische sowie syntaktische Integrität der Texte zu wahren. Ein zentrales Bereinigungsziel bestand in der Erhaltung der Satzstruktur, um eine präzise Analyse der Lesbarkeit zu gewährleisten.

3.4.2.1 Datenaufbereitung

Zunächst wurden mithilfe eines hierfür eigens entwickelten Python-Codes Tabellen, Abbildungen, Überschriften, Links sowie Aufzählungen ohne abschließende Satzzeichen entfernt, um den Fließtext der Berichte zu erfassen.[74] Um sicherzustellen, dass Punkte ausschließlich als Satzendzeichen identifiziert wurden, wurden insgesamt 209 spezifische sowie gängige Abkürzungen berücksichtigt. Diese wurden im Rahmen der Bereinigung durch ihre ausgeschriebene Fassung ersetzt. Zwecks Verbesserung der maschinellen Satzsegmentierung wurden Punkte innerhalb von Sätzen – bspw. hinter Datumsangaben –eliminiert.[75] Die Konvertierung in Word-Dokumente führte vereinzelt zu Umbrüchen innerhalb von Wörtern. Diese Unschärfe wurde korrigiert, indem die getrennten Wörter wieder zusammengeführt wurden. Ferner wurden Bindestriche durch Leerzeichen ersetzt, Apostrophe sowie entsprechende HTML-Zeichencodes entfernt und mehrfach aufeinanderfolgende Leerzeichen auf ein einzelnes reduziert. Zusätzlich wurden nicht-alphabetische Zeichen, mit Ausnahme von Umlauten und relevanten Satzzeichen, durch Leerzeichen ersetzt. Aufeinanderfolgende Punkte wurden zu einem Punkt zusammengeführt. Abschließend wurden im Rahmen der Lesbarkeitsalgorithmen überflüssige Leerzeichen eliminiert. Die maschinelle Bereinigung wurde iterativ verbessert, indem Besonderheiten einzelner Vergütungsberichte in den Python-Code sukzessiv eingearbeitet wurden.[76]

[73] Siehe Kapitel 3.5

[74] Ein ähnliches Vorgehen diesbezüglich findet sich bei *Laksmana et al.* (2012); im Gegensatz dazu schließen *Quick et al.* (2024) sämtliche Fußnoten aus, die in dieser Untersuchung bei vorliegender Satzstruktur erhalten blieben.

[75] Vgl. auch Vorgehen bei *Quick et al.* (2024).

[76] Die angewandte Bereinigungsmethodik lehnt sich hauptsächlich an das Verfahren von *Hemmings et al.* (2020, S. 584 ff.) an. Während die Autoren nur Tabellen o. Ä. entfernt haben, deren Inhalt zu mehr als 15 % aus Zahlen bestand, wurden in der vorliegenden Untersuchung Untersuchung sämtliche Tabellen und Grafiken unabhängig von ihrem narrativen Textanteil entfernt (vgl. Kapitel 3.6 Limitation).

Die maschinell bereinigten Berichte (n=426)[77] wurden abschließend einer umfassenden manuellen Prüfung unterzogen. Die Einbindung manueller Prüfungen erwies sich als gängige Methode, um Formatierungsabweichungen und spezifische Eigenheiten einzelner Dokumente zu adressieren.[78] Besonders bei Berichten mit komplexer Formatierung war eine manuelle Nachbearbeitung unerlässlich. Derartige Berichte enthielten bspw. bis zu fünf aufeinanderfolgende Absätze, Unterbrechungen durch grafische Elemente, nicht-standardisierte Textstrukturen oder lagen lediglich als gescannte Druckversion vor.

3.4.2.2 Maschinelle Analyse und Validierung der Ergebnisse

Für die Lesbarkeitsanalyse der Vergütungsberichte wurden zwei Python-Codes entwickelt, die auf die Anforderungen der deutschen und englischen Sprache zugeschnitten sind. Dies war notwendig, da sprachspezifische Unterschiede, insbesondere bei der Silbenerkennung, eine getrennte Verarbeitung erforderten. Während der deutsche Code unter Verwendung der *spaCy*-Bibliothek das Modell *de_core_news_sm* nutzt, greift der englische Code auf das Modell *en_core_web_sm* zurück. Das allgemeine Vorgehen beider Codes umfasst die präzise Erkennung von Sätzen, Wörtern und Silben sowie die Berechnung der oben eingeführten Lesbarkeitsindizes.

Die Entscheidung zur Entwicklung eines eigenen Codes anstelle der Nutzung verfügbarer Online-Tools wurde getroffen,[79] um methodische Schwächen externer Lösungen zu vermeiden. Stichprobenartige Validierungstests mit drei webbasierten Tools wiesen inkonsistente Ergebnisse auf, darunter Fehler bei der Satzsegmentierung und eine fehlende Transparenz hinsichtlich der Berechnungsprozesse. Der entwickelte Code adressiert diese Herausforderungen durch die detaillierte Darstellung aller Zwischenschritte und ermöglicht so eine präzise sowie nachvollziehbare Analyse durch die Ausgabe erkannter Wörter, Sätze oder die Silbenzählung pro Wort sowie weiterer Zwischenergebnisse in Excel. Die getesteten Online-Tools[80] wurden ergänzend als Robustheitstest eingesetzt, um den entwickelten Code neben stichprobenbasierten Prüfungen der eigenen Zwischenergebnisse zu validieren. Die Ergebnisse von bis zu drei verschiedenen Tools wurden dabei

[77] Von den insgesamt 438 erhobenen Berichten wurden 426 in die Lesbarkeitsanalyse einbezogen. Aufgrund technischer Restriktionen konnten die Berichte dreier Unternehmen dahingehend nicht analysiert werden, etwa weil es sich um gescannte Dokumente handelte, bei denen die OCR-Texterkennung fehlschlug. Durch die Analyse der verbleibenden Berichte entstanden über zwei Millionen Datenpunkte (bspw. Zwischenergebnisse der Lesbarkeitsanalyse).
[78] Den Verfassern sind keine Studien bekannt, die eine vergleichbare manuelle Bereinigungsschleife vorsehen oder explizit darauf eingehen.
[79] Andere Autoren nutzen frei verfügbare Webtools. Vgl. bspw. *Diederich/Velte* (2019), S. 465; *Heilmann et al.* (2020), S. 711 sowie *Quick et al.* (2023), S. 848.
[80] *Online Utility* (2025); *Readability Formulas* (2024); *Schöll* (2025); *Bond* (2025); *Lenhard/Lenhard* (2022).

gemittelt und die Differenz zu dem eigenen Code berechneten Werten analysiert. Diese Differenzen waren durchgehend marginal und beeinflussten die Einstufung der Lesbarkeit nicht.

3.5 Ergebnisse der Untersuchung und Diskussion

3.5.1 Ort der Veröffentlichung

Ungeachtet der Veröffentlichung auf der Internetseite stellt sich für Unternehmen die praktische Frage, den Vergütungsbericht in die konventionelle Berichterstattung einzubinden.[81] Ein explizites Berichtsformat bzw. der Standort des Vergütungsberichts wird in § 162 AktG nicht genannt. Somit ist grundsätzlich eine Veröffentlichung außerhalb oder innerhalb des Geschäftsberichts denkbar.[82]

Mit 43 der 73 untersuchten Unternehmen haben die Unternehmen mehrheitlich den Vergütungsbericht in den Geschäftsbericht integriert. Dies entspricht sowohl für 2021 als auch 2022 einem Anteil von 58,90 %. Die restlichen Unternehmen der Stichprobe haben den Vergütungsbericht außerhalb des Geschäftsberichts und somit als ein ausschließlich separates Berichtsinstrument auf der Unternehmenshomepage als PDF-Datei veröffentlicht. Auffällig ist, dass der Vergütungsbericht in acht Fällen für das Berichtsjahr 2021 und in fünf Fällen für das Jahr 2022 in den Lagebericht integriert wurde. Dies widerspricht der verbreiteten Rechtsauffassung, wonach der Vergütungsbericht als aktienrechtliches Informationsinstrument zwingend getrennt vom Lagebericht als handelsrechtliches Informationsinstrument zu veröffentlichen ist.[83] Diese Trennung wird insbesondere mit der unterschiedlichen Zuständigkeitsverteilung begründet: Während der Vergütungsbericht als Gemeinschaftsaufgabe von Aufsichtsrat und Vorstand angefertigt wird, obliegt die Erstellung des Lageberichts gemäß § 264 Abs. 1 HGB ausschließlich dem Vorstand.[84]

Unabhängig des gewählten Standorts ist der Vergütungsbericht gemäß § 162 Abs. 4 AktG pflichtmäßig auf der Internetseite zu veröffentlichen und kostenfrei zugänglich zu machen. Allerdings zeigt sich, dass die Auffindbarkeit der (archivierten) Vergütungsberichte nicht in jedem Fall ohne hohe Suchkosten möglich ist. Zwar gilt die Publizitätspflicht nach herrschender Meinung bereits mit der Bekanntmachung der Hauptversammlungsunterlagen als erfüllt.[85] Dem steht jedoch entgegen, dass – auch wenn die Aktionäre aus betriebswirtschaftlicher Sicht als primäre Adressaten der

[81] Vgl. *Vetter/Tielmann* (2022), S. 387; *Grigoleit/Rachlitz*, AktG § 162 Rn. 73.
[82] Da der Vergütungsbericht vor ARUG II kein eigenständiges Berichtsinstrument und Bestandteil des Geschäftsberichts war, werden im Folgenden nur die Berichtsjahre 2021 und 2022 betrachtet.
[83] Vgl. *Vetter/Tielmann* (2022), S. 387; *Kuthe/Reuters* (2022), S. 2116.
[84] Vgl. *Vetter/Tielmann* (2022), S. 387; *MüKoAktG/Spindler*, AktG § 162 Rn. 11a f.
[85] Vgl. *MüKoAktG/Spindler*, AktG § 162 Rn. 91; *Grigoleit/Rachlitz*, AktG § 162 Rn. 72; siehe auch Gesetzesbegründung: *Deutscher Bundestag* (2019), S. 113.

Vergütungsberichterstattung gelten – ein erweiterter Adressatenkreis zu berücksichtigen ist. So betont etwa der Erwägungsgrund 33 der Europäischen Aktionärsrichtlinie (Richtlinie (EU) 2017/828), dass Offenlegungspflichten auch potenziellen Anlegern und an der Gesellschaft interessierten Akteuren eine Beurteilung der Vergütung ermöglichen sollen. Vor diesem Hintergrund erscheint eine ausschließlich aktionärsorientierte Veröffentlichung im Kontext des heutigen Corporate-Governance-Reportings nicht mehr zweckmäßig. Es wird daher empfohlen, Vergütungsberichte auf der Unternehmenswebsite klar gekennzeichnet und über einen einheitlichen Menüpunkt sowie eine verständliche Menüführung – etwa im Bereich Investor Relations (IR) – leicht zugänglich bereitzustellen.[86]

Bei 54 der 73 Unternehmen sind die Vergütungsberichte der Berichtsjahre 2021 sowie 2022 auf den Investor-Relations-Seiten klar gekennzeichnet, chronologisch aufgeführt und somit problemlos zugänglich. Des Weiteren beginnt bei einem Unternehmen die chronologische Auflistung erst mit dem Berichtsjahr 2022. Vier Unternehmen bieten die Berichte weder explizit auf den Investor-Relations-Seiten noch dauerhaft in den Unterlagen zur Hauptversammlung an, sodass der Zugang ausschließlich über den Geschäftsbericht möglich ist. Bei 14 Unternehmen sind die Vergütungsberichte nur über die archivierten Unterlagen in den verschiedenen Ordnern der Hauptversammlungen möglich gewesen. Bei der Erhebung hat sich gezeigt, dass einige Unternehmen lediglich den Vergütungsbericht der letzten Berichtsperiode auf den Corporate-Governance-Seiten ihrer Website gut sichtbar platzieren.[87] Während die Geschäftsberichte in aller Regel im Archiv der Finanzpublikation chronologisch gelistet werden, ist der Zugang der nicht im Geschäftsbericht eingebetteten Vergütungsberichte somit nur über die archivierten offengelegten Unterlagen der jeweiligen Hauptversammlungen möglich gewesen. Es ist fraglich, ob diese Publikationspraxis dem in Erwägungsgrund 32 der Europäischen Aktionärsrichtlinie (Richtlinie (EU) 2017/828) geforderten „problemlosen Zugang“ tatsächlich entspricht.

3.5.2 Umfang, Abbildungen und weitere formale Gestaltung

In *Tabelle 3.1* sowie *Abbildung 3.2* werden die Ergebnisse hinsichtlich des Umfangs sowie der formalen Gestaltung der Vergütungsberichte präsentiert. Gemessen anhand der Anzahl der Wörter zeigt sich, dass der Berichtsumfang der Vergütungsberichte 2021 nach Einführung des ARUG II

[86] Vgl. *BeckOGK/Bayer/Scholz*, AktG § 162 Rn. 169. Im Rahmen der Untersuchung wurde die leichte Zugänglichkeit der Vergütungsberichte angenommen, sofern diese für jedes Berichtsjahr klar gekennzeichnet waren und entweder als eigener Menüpunkt auf den IR-Seiten aufgeführt wurden oder über Menüpunkte wie „Vergütung“, „Corporate Governance“, „Vorstand“ oder „Publikationen“ ohne wesentliche Hürden zugänglich waren.

[87] Zum Zeitpunkt der Erhebung betraf dies Vergütungsberichte des Jahres 2023, die nicht Gegenstand der Untersuchung sind.

signifikant gestiegen ist ($p < 0{,}0001$). Die durchschnittliche Wortanzahl der Berichte stieg von 4043,38 (2020) auf 6494,18 (2021). Im Berichtsjahr 2022 beträgt die durchschnittliche Anzahl der Wörter 6666,44, womit der Berichtsumfang erneut gegenüber dem Vorjahr angestiegen ist. Jedoch lässt sich anhand der Standardabweichung feststellen, dass der Berichtsumfang der Unternehmen über die Organvergütung teils erheblich variiert. So umfasst der Vergütungsbericht der TAG Immobilien AG im Jahr 2022 lediglich 2536 Wörter. Hingegen sind die Vergütungsberichte der Deutschen Bank AG über alle drei Jahre mit Abstand am umfangreichsten (2020: 12639 Wörter; 2021: 15635 Wörter; 2022: 14627 Wörter). Dies ist dadurch begründet, dass die Deutsche Bank AG zusätzlich über die Vergütung der Mitarbeitenden, die nicht im Vorstand oder Aufsichtsrat sitzen, Bericht erstattet. Das Unternehmen kommt damit zusätzlichen gesetzlichen Anforderungen aus der Institutsvergütungsverordnung sowie Kapitaladäquanzverordnung nach.

	Deutsch			Englisch		
	2020	2021	2022	2020	2021	2022
Mittelwert	4043,38	6494,18	6666,44	5204,63	7703,87	8266,85
Median	3938	6105	6116	5185	7184	7928
Standardabweichung	2165,30	2490,39	2286,41	2667,09	2822,25	2775,98
Stichprobenumfang[88]	71	71	71	71	71	71

Tabelle 3.1: Anzahl der Wörter in den Vergütungsberichten

Ähnliche Entwicklungen lassen sich auch in den in englischer Sprache verfassten Vergütungsberichten beobachten. Die durchschnittliche Wortanzahl dieser Berichte ist ebenfalls signifikant angestiegen ($p < 0{,}0001$). In der ersten Berichtsperiode nach den Vorgaben des ARUG II umfassen die Vergütungsberichte durchschnittlich 7703,87 Wörter, was einer durchschnittlichen Zunahme von 2499,24 Wörtern pro Bericht im Vergleich zur letzten Berichtsperiode vor Inkrafttreten des ARUG II entspricht. Ein marginaler Anstieg des Berichtsumfangs zwischen den Jahren 2021 und 2022 sowie die Variationen im Umfang der Berichte zeigen sich analog zu den Entwicklungen bei den Vergütungsberichten in deutscher Sprache.

Der starke Anstieg der Anzahl an Wörtern deutet an, dass die Berichterstattung umfangreicher und möglicherweise komplexer geworden ist. Diese Entwicklung kann im Kontext der *Information*

[88] Die Ermittlung der Anzahl der Wörter wurde anhand der Datengrundlage für die folgende Lesbarkeitsanalyse durchgeführt. Siehe zum Vorgehen Kapitel 3.2.

Overload-Theorie gesehen werden. Information Overload wird in der Accounting-Forschung als Zustand definiert, in dem die Anforderungen an die Informationsverarbeitung einer Person deren verfügbare Zeit oder Kapazität zur Verarbeitung aufgrund zu vieler oder irrelevanter Informationen übersteigen.[89] Um Information Overload zu begegnen, kann der Einsatz von Visualisierungen (z. B. Abbildungen und Tabellen) dienlich sein, um Informationen zu geringen Kosten zu vermitteln und die Kapazität des Adressaten zur Informationsverarbeitung zu entlasten.[90] Eine Einbettung von Abbildungen und Tabellen wird auch in der Regierungsbegründung angeregt.[91] Ebenfalls kann der Einsatz von Inhaltsverzeichnissen die Suchzeit des Adressaten nach gezielten Informationen reduzieren und den Bericht besser strukturieren. Angesichts dessen werden im Folgenden die Ergebnisse hinsichtlich der Einbettung von Abbildungen und Inhaltsverzeichnissen wiedergegeben. *Abbildung 3.2* zeigt die Entwicklung der Gestaltung von Vergütungsberichten anhand der Nutzung von Abbildungen und Tabellen. Zudem wird der durchschnittliche Seitenumfang der Vergütungsberichte im Zeitverlauf dargestellt, um das Verhältnis zwischen visuellen Darstellungen und dem Seitenumfang zu analysieren.

Die gestiegene Anzahl an Wörtern zeigt sich auch im Seitenumfang der Berichte. Während der durchschnittliche Vergütungsbericht aller 73 untersuchten Unternehmen noch 15,1 Seiten (Median: 14 Seiten) im Jahr 2020 umfasste, ist die durchschnittliche Seitenzahl auf 27,75 (2021 – Median: 26 Seiten) bzw. 30,11 (2022 – Median: 28 Seiten) gestiegen. Dies ist ein signifikanter Anstieg ($p < 0{,}0001$) zwischen den ersten beiden Jahren, wenngleich sich die teils erhebliche Streuung der Anzahl an Wörtern auf den Seitenumfang übertragen lässt. Während die durchschnittliche Anzahl an Abbildungen und Tabellen pro Vergütungsbericht in den Jahren sukzessive von 11,52 Abbildungen bzw. Tabellen (2020 – Median: 10) auf 19,64 (2021 – Median: 19) bzw. 21,05 (2022 – Median: 21) gestiegen ist und somit einen signifikanten Anstieg ($p < 0{,}0001$) mit erstmaliger ARUG-II-Anwendung verzeichnet, hat die inhaltliche Gewichtung der Abbildungen und Tabellen im Verhältnis zur Gesamtlänge der Berichte abgenommen. Der relative Anteil der Abbildungen bzw. Tabellen je Seite ist kontinuierlich gesunken: Im Jahr 2020 lag dieser Wert bei 0,77, reduzierte sich 2021 auf 0,74

[89] Vgl. *Schick et al.* (1990), S. 199; *Hartmann/Weißenberger* (2024), S. 1622.
[90] Vgl. hierzu die Anmerkungen des *IASB* (2017, S. 24) zu den Lösungsansätzen des sogenannten „Disclosure Problem". Als „Disclosure Problem" wird die unzureichende Kommunikation relevanter Informationen sowie die geringe Berücksichtigung wichtiger Informationen und die Einbettung von zu vielen irrelevanten Informationen in Jahresabschlüssen bezeichnet.
[91] Vgl. *Deutscher Bundestag* (2019), S. 109. Jedoch kritisieren Aktionärsverbände dieses Vorgehen vereinzelt. So bemängelt die Schutzgemeinschaft der Kapitalanleger (SdK) in ihrer Beschlussempfehlung zum Vergütungsbericht 2023 der MorphoSys AG die „*eng geschriebenen Seiten mit zahlreichen Tabellen und Grafiken*" als zu komplex. Vgl. hierzu *Schutzgemeinschaft der Kapitalanleger* (2024).

und erreichte 2022 einen Wert von 0,72. Es kann festgestellt werden, dass dem gestiegenen Komplexitätsbedarf des Berichts gemessen am Umfang nicht im gleichen Maße mit einer gestiegenen Unterstützung der Informationen durch Tabellen und Abbildungen begegnet wurde.

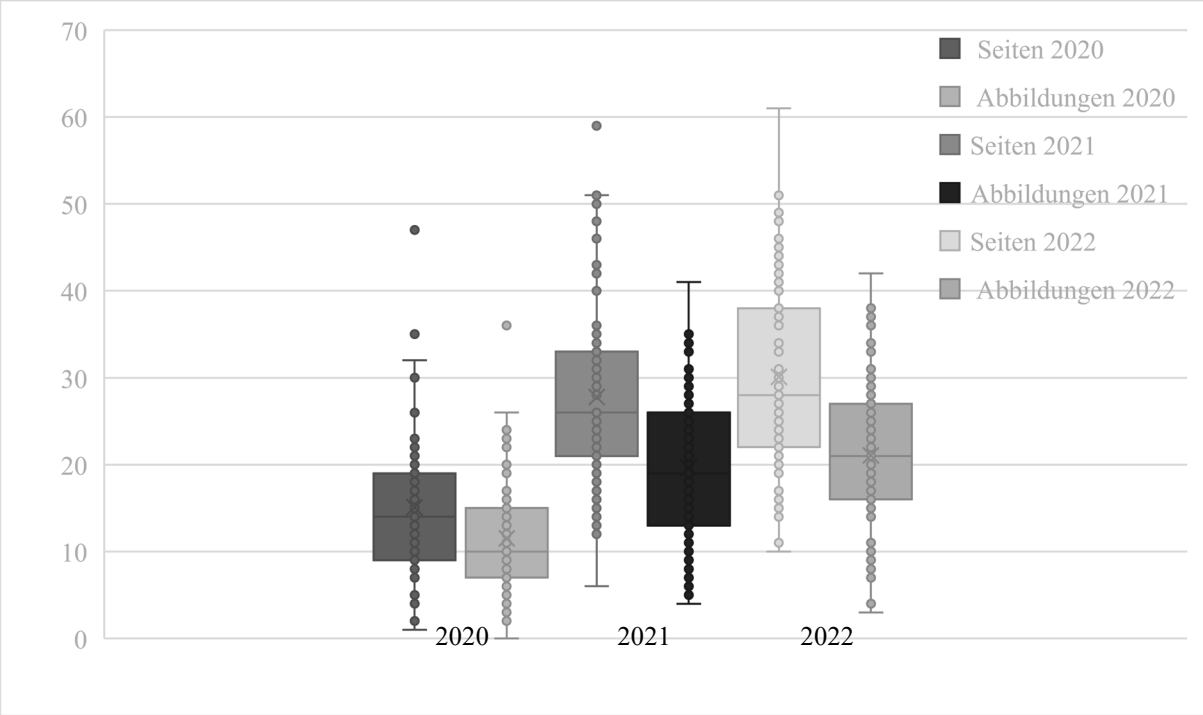


Abbildung 3.2: Umfang und Gestaltung der Berichte

Im Jahr 2021 enthielten 27 Vergütungsberichte (36,99 % der untersuchten Unternehmen) ein Inhaltsverzeichnis, während dieser Anteil im darauffolgenden Berichtsjahr auf 36 Vergütungsberichte (49,32 %) anstieg.

3.5.3 Prüfung der Berichte 2021 und 2022

Gemäß § 162 Abs. 3 AktG ist der Vergütungsbericht vom Abschlussprüfer dahingehend zu prüfen, ob die erforderlichen Mindestangaben im Vergütungsbericht gemäß § 162 Abs. 1 und 2 AktG gemacht wurden. Die Veröffentlichungspflicht nach § 162 Abs. 4 AktG umfasst auch die Offenlegung des Prüfvermerks des Abschlussprüfers. Für das Berichtsjahr 2021 sind alle 73 untersuchten Unternehmen dieser Pflicht nachgekommen und haben den Prüfvermerk veröffentlicht. Im Berichtsjahr 2022 fehlt der Prüfvermerk jedoch vollständig bei den Vergütungsberichten der Aixtron SE und Evotec SE. Beide Unternehmen haben ihren Vergütungsbericht außerhalb des Konzernabschlusses und des Konzernlageberichts veröffentlicht, sodass der Bestätigungsvermerk des Abschlussprüfers keinen Bezug zum Vergütungsbericht nimmt und daher nicht als Ersatztestat gewertet werden kann. In diesen Fällen fehlt den externen Adressaten somit der Prüfvermerk als Form der Assurance. Hinsichtlich der

Prüfungstiefe haben sowohl 2021 als auch 2022 insgesamt 44 Unternehmen eine inhaltliche Prüfung ihrer Vergütungsberichte durchgeführt, was etwa 60 % der untersuchten Unternehmen entspricht. Auffällig ist, dass sechs Unternehmen von einer formellen Prüfung 2021 auf eine inhaltliche Prüfung 2022 umgestellt haben. Im gleichen Umfang haben jedoch sechs Unternehmen ihre Prüfungstiefe reduziert und von einer inhaltlichen Prüfung 2021 auf eine formelle Prüfung 2022 gewechselt. Gründe für diese Änderungen werden nicht angegeben. Branchenspezifische Besonderheiten zeigen sich in der Automobil- und Chemiebranche, in der alle Unternehmen ihren Vergütungsbericht 2022 freiwillig einer inhaltlichen Prüfung unterzogen haben. Ähnlich wie bei der Überprüfung der Angemessenheit der Vorstandsvergütung, bei der eine Analyse vergleichbarer Unternehmen derselben Branche erfolgt, könnte auch das Berichtsverhalten branchengleicher Unternehmen als Maßstab für die Prüfungstiefe dienen.

Eine inhaltliche Prüfung deckt die gesetzlich vorgeschriebene formelle Pflichtprüfung nach § 162 Abs. 3 AktG ab. Im Sinne einer guten Corporate Governance ist es daher aus Adressatensicht begrüßenswert, wenn die Vergütungsberichte neben der pflichtmäßigen formellen Prüfung auch materiell durch den Abschlussprüfer hinsichtlich der inhaltlichen Richtigkeit geprüft werden. Die vom IDW geäußerten Bedenken hinsichtlich einer möglichen vergrößerten Erwartungslücke durch eine lediglich formelle Prüfung sind unter Berücksichtigung der Analyseergebnisse somit unbegründet.[92] Es bleibt allerdings abzuwarten, ob zukünftig weitere Unternehmen auf eine materiell-inhaltliche Prüfung setzen werden.

3.5.4 Lesbarkeit

Die bisherigen Ergebnisse können als Indikator einer erhöhten Komplexität interpretiert werden. Diese Schlussfolgerung stünde im Gegensatz zur Formulierung des § 162 Abs. 1 Satz 1 AktG, der einen klaren und verständlichen Vergütungsbericht vorschreibt. Mittels Berechnung von Lesbarkeitsindizes ist eine differenzierte Beurteilung der Vergütungsberichte möglich. Im Folgenden werden in *Tabelle* 3.2 sowie *Abbildung 3.3* und *Tabelle 3.3* und *Abbildung 3.4* die Ergebnisse der ermittelten Lesbarkeitswerte der Vergütungsberichte in deutscher und englischer Sprache im Zeitraum 2020 bis 2022 dargestellt und anschließend eingeordnet.

Die Lesbarkeit der englischsprachigen Vergütungsberichte wird anhand der in *Kapitel 3.3.3* vorgestellten Lesbarkeitsindizes FRE sowie FOG erhoben. Der durchschnittliche FRE-Wert ist von 25,24 im Jahr 2020 auf 22,60 im Jahr 2021 gesunken. Die Veränderung im Prä-Post-Vergleich zur

[92] Vgl. *IDW* (2018), S. 4.

ARUG-II-Einführung ist statistisch signifikant ($p < 0{,}0001$) und deutete auf eine erschwerte Lesbarkeit der Vergütungsberichte hin. Im Jahr 2022 ist der Mittelwert des FRE marginal auf 22,44 gesunken. Entsprechend der Klassifizierung des FRE wird der Text somit konstant als sehr schwierig eingestuft und richtet sich primär an ein Publikum mit akademischem Hintergrund.[93] Eine ähnliche Entwicklung lässt sich für den FOG beobachten. Ausgehend von einem FOG-Wert von 20,26 im Jahr 2020 stieg der Mittelwert statistisch signifikant ($p < 0{,}0001$) auf 21,00 im Jahr 2021 an und veränderte sich im Jahr 2022 nur marginal auf 21,05. Nach den theoretischen Referenzwerten des FOG entsprechen die Vergütungsberichte über den Beobachtungszeitraum somit dem Leselevel von Hochschulabsolventen.[94]

	Flesch Reading Ease (FRE)			Gunning Fog Index (FOG)		
	2020	2021	2022	2020	2021	2022
Mittelwert	25,24	22,60	22,44	20,26	21,00	21,05
Minimum	13,04	13,41	14,71	17,39	18,80	19,39
Maximum	34,58	29,30	29,49	23,77	23,76	23,63
Standardabweichung	3,90	3,33	3,06	1,22	0,96	0,87
Stichprobenumfang[95]	71	71	71	71	71	71

Tabelle 3.2: Die Lesbarkeit der englischsprachigen Vergütungsberichte

Des Weiteren variieren die Lesbarkeitswerte mit der Branchenzugehörigkeit. In dem untersuchten Sample sind die Vergütungsberichte der Unternehmen aus dem Energiesektor für die Berichtsjahre 2021 und 2022 mit einem durchschnittlichen FOG-Wert von 19,76 am einfachsten zu lesen, während die Vergütungsberichte der Unternehmen aus der Medienbranche am komplexesten sind.

Beide verwendeten Lesbarkeitsindizes deuten einen gleichartigen Trend an, so dass robuste Ergebnisse geschlussfolgert werden können. Grundsätzlich lässt sich eine marginale Verschlechterung der Lesbarkeit der englischsprachigen Vergütungsberichte mit ARUG II feststellen. Insgesamt liegen die Mittelwerte des FRE und FOG für die englischsprachigen Berichte in einem Bereich, der auf eine sehr schwere Lesbarkeit bzw. Verständlichkeit hinweist.

[93] *Vgl. Kapitel 3.3.3.*
[94] *Vgl. Kapitel 3.3.3.*
[95] Diese sowie fortfolgende Analysen wurden anhand der Datengrundlage für die Lesbarkeitsanalyse durchgeführt. Zum Vorgehen siehe *Kapitel 3.4.2.*

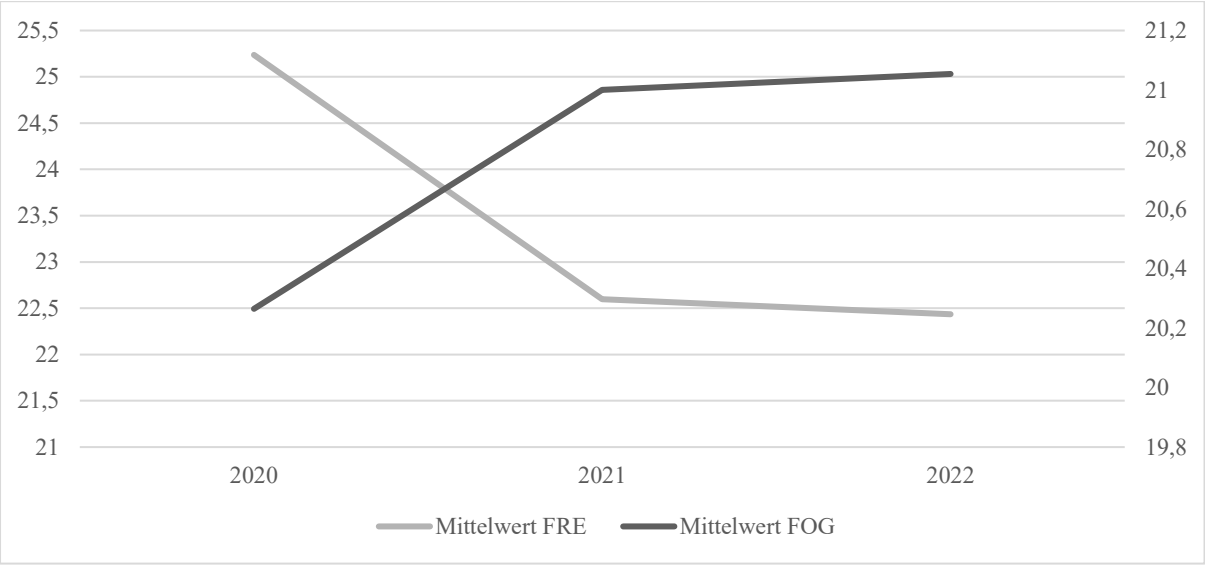


Abbildung 3.3: Entwicklung der Mittelwerte der Lesbarkeitsindizes FRE und FOG

Die Ergebnisse des FOG zeigen für die untersuchten Vergütungsberichte einen Mittelwert von 20,77 über den Zeitraum 2020 bis 2022. Dieser Wert liegt deutlich über den empfohlenen Schwellenwerten für geschäftliche Texte von maximal zwölf.[96] Während die Ergebnisse auf einem ähnlichen Niveau wie Studien zu Vergütungsberichten von US-Unternehmen liegen,[97] deutet eine Untersuchung der Vergütungsberichte von FTSE 350-gelisteten Unternehmen auf eine bessere Lesbarkeit hin.[98] Amerikanische Vergütungsberichte, die den regulatorischen Anforderungen der SEC unterliegen, weisen einen durchschnittlichen FOG von 21,82 auf.[99] Dieser Wert liegt leicht über den Befunden der vorliegenden Analyse und unterstreicht, dass auch amerikanische Vergütungsberichte ein hohes Maß an sprachlicher Komplexität aufweisen. Dabei ist jedoch zu berücksichtigen, dass sich die regulatorischen Rahmenbedingungen in den USA deutlich von denen in Deutschland unterscheiden.

Weitere Untersuchungen bieten Kontext zur Textkomplexität anderer Berichtsinstrumente. In einer Studie zu Pflichtangaben nach IFRS 3 wurde ein FOG-Wert von 19,05 ermittelt, wobei geografische und sprachliche Überschneidungen mit dem vorliegenden Sample bestehen.[100] Für Key Audit Matters (KAMs) wurden FOG-Werte von 22,51 für das Jahr 2020 und 22,36 für 2021 ermittelt, die nahe an den Ergebnissen der vorliegenden Untersuchung liegen.[101] Obwohl sich KAMs, ähnlich wie Vergütungsberichte, an Aktionäre richten, handelt es sich um Kommunikationselemente, die von

[96] Vgl. *Laksmana et al.* (2012), S. 191.
[97] Vgl. *Hemmings et al*. (2020), S. 557: Mittelwert 23,6; *Laksmana et al.* (2012), S. 191: Mittelwert 21,82.
[98] Vgl. *Hooghiemstra et al.* (2017), S. 706: Mittelwert > 17.
[99] Vgl. *Laksmana et al.* (2012), S. 191.
[100] Vgl. *Rupertus et al.* (2017), S. 83.
[101] Vgl. *Quick et al.* (2023), S. 851.

Abschlussprüfern bereitgestellt werden. Untersuchungen zur Lesbarkeit der lageberichtähnlichen Berichterstattung mittels MD&A-Berichterstattung US-amerikanischer Unternehmen (FOG: 18,02 bzw. 18-20)[102] verdeutlichen, dass Vergütungsberichte deutscher Unternehmen sprachlich anspruchsvoller gestaltet sind.

Die englischsprachigen Vergütungsberichte der HDAX-Unternehmen weisen im Durchschnitt einen FRE-Mittelwert auf, der zwischen den Ergebnissen von *Laksmana et al.*[103] und *Velte* liegt. Ein Vergleich mit der Studie von *Velte*, der Vergütungsberichte ausgewählter HDAX-Unternehmen für den Zeitraum 2014-2018 analysierte, zeigt einen FRE-Mittelwert von 26,64. Dieser Wert liegt zwar über dem Ergebnis der vorliegenden Untersuchung, fällt jedoch weiterhin in die anspruchsvollste Lesbarkeitskategorie (0-30).[104]

Untersuchungen zu anderen Berichtsinstrumenten zeigen ebenfalls eine hohe sprachliche Komplexität, weisen jedoch Unterschiede in Zielsetzung, Struktur und Lesbarkeit im Vergleich zu Vergütungsberichten auf.[105]

Die Lesbarkeit der deutschsprachigen Vergütungsberichte wird anhand der in *Kapitel 3.3.3* vorgestellten Indizies LIX sowie FRE_{Amstad} erhoben. Die Ergebnisse sind in *Tabelle 3.3* und *Abbildung 3.4* dargestellt. Der durchschnittliche LIX-Wert der Vergütungsberichte im Jahr 2020, der letzten Berichtsperiode vor der Anwendung des ARUG II, beträgt 65,39. Im Zuge der Implementierung des ARUG II stieg der LIX-Wert der Vergütungsberichte im Prä-Post-Vergleich statistisch signifikant ($p < 0{,}0001$) auf durchschnittlich 67,44 im Jahr 2021 an, was auf eine erschwerte Lesbarkeit der Berichte hinweist. In der zweiten Berichtsperiode sank der durchschnittliche LIX-Wert im Jahr 2022 marginal auf 67,29. Die Vergütungsberichte liegen stets über dem Schwellenwert von 60 und sind demgemäß als sehr schwer verständlich zu klassifizieren.[106]

[102] Vgl. *Lo et al.* (2017), S. 7.
[103] Vgl. *Laksmana et al.* (2012), S. 191: Mittelwert 19,84.
[104] Vgl. *Velte* (2021), S. 38.
[105] *Kaya/Seebeck* (2018, S. 1000) analysieren KAMs und ermittelten FRE-Werte von 20,14 bis 30,29; *Rupertus et al.* (2017, S. 83) errechnen für Finanzinformationen (z. B. IFRS 3) einen FRE-Wert von 22,81.
[106] *Vgl. Kapitel 3.3.3.*

	Lesbarkeitsindex nach Björnsson (LIX)			Flesch Reading Ease Amstad (FRE_{Amstad})		
	2020	2021	2022	2020	2021	2022
Mittelwert	65,39	67,44	67,29	25,77	23,22	23,66
Minimum	58,17	62,63	61,93	12,10	15,04	14,94
Maximum	74,41	74,19	74,39	34,57	29,09	30,78
Standardabweichung	2,96	2,29	2,20	3,89	3,19	2,75
Stichprobenumfang	71	71	71	71	71	71

Tabelle 3.3:Die Lesbarkeit der deutschsprachigen Vergütungsberichte

Eine vergleichbare Entwicklung der Lesbarkeit lässt sich anhand des FRE_{Amstad} beobachten. Ausgehend von einem durchschnittlichen Lesbarkeitswert von 25,77 nach FRE_{Amstad} im Jahr 2020, zeigte sich mit der erstmaligen Anwendung der neuen gesetzlichen Vorgaben ein statistisch signifikanter Rückgang ($p < 0{,}0001$) auf 23,22 im Jahr 2021, was ebenfalls auf eine erschwerte Lesbarkeit hindeutet. Analog zur Entwicklung des LIX-Werts veränderte sich der FRE_{Amstad}-Wert im Jahr 2022 nur marginal und erreichte einen durchschnittlichen Wert von 23,66. Die Lesbarkeitswerte nach FRE_{Amstad} liegen durchweg in der Referenzkategorie von 20 bis 30 und weisen auf eine schwierige Lesbarkeit der Vergütungsberichte hin.[107] Die Ergebnisse lassen auf robuste Ergebnisse schließen. Besonders hervorzuheben sind die Vergütungsberichte der Deutschen Post AG (durchschnittlicher LIX-Wert: 62,95) sowie der RWE AG (durchschnittlicher LIX-Wert: 63,64), die in der Post-ARUG-II-Berichtsperiode 2021/2022 die niedrigsten LIX-Werte aufweisen und damit im Vergleich zu anderen Berichten eine relativ gute Lesbarkeit aufzeigen.[108] Im Branchenvergleich erzielen Unternehmen aus der Energiebranche sowie dem Bankensektor mit einem durchschnittlichen LIX-Wert von 64,26 bzw. 64,69 die besten Lesbarkeitswerte.

[107] *Vgl. Kapitel 3.3.3.*
[108] Diese Feststellung lässt sich ebenfalls auf die englischsprachigen Vergütungsberichte der Jahre 2021 bzw. 2022 übertragen (ermittelt über den FOG-Wert).

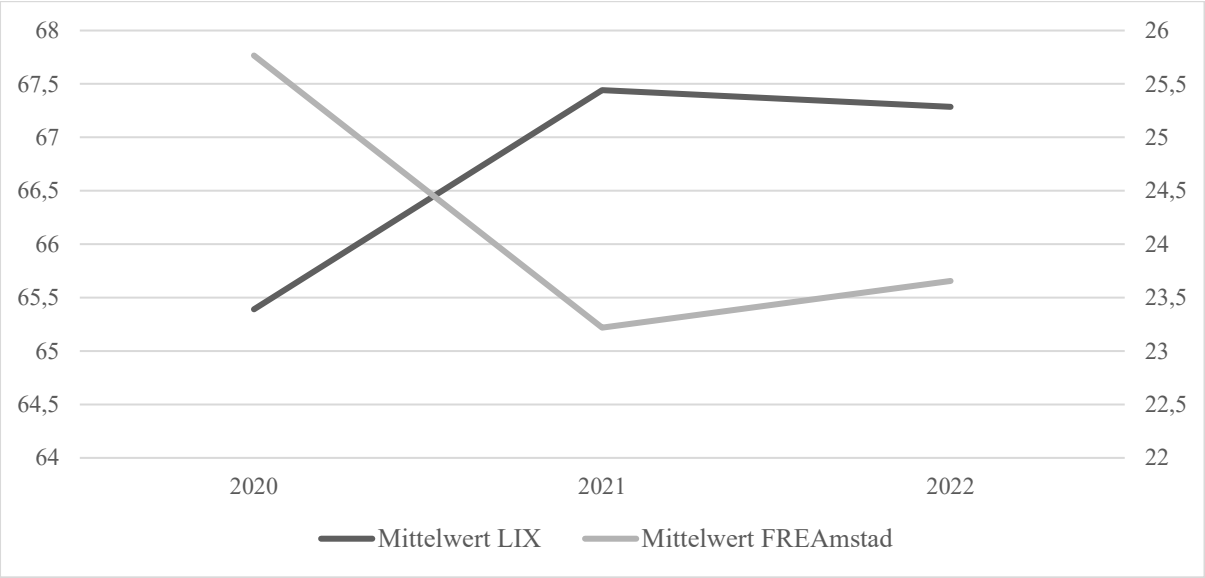


Abbildung 3.4: Entwicklung der Mittelwerte der Lesbarkeitsindizes LIX und FRE_{Amstad}

Ein direkter Vergleich der ermittelten Ergebnisse für deutsche Vergütungsberichte ist schwierig, da bisher keine textanalytischen Untersuchungen bekannt sind, die deutschsprachige Vergütungsberichte mit einem vergleichbaren Sample und unter Verwendung der Lesbarkeitsindizes LIX und FRE_{Amstad} analysiert haben. Mit einem LIX-Mittelwert von 66,70 und einem FRE_{Amstad}-Mittelwert von 24,21 können unsere Ergebnisse lediglich als indirekter Robustheitstest für die englischen Analysen herangezogen werden, gleichwohl die Indizes sprachübergreifend nicht vergleichbar sind. Das Fehlen solcher Studien unterstreicht hingegen die Relevanz unserer Untersuchung im deutschsprachigen Raum. Vergleichbare Studien zu anderen Berichtsinstrumenten liefern jedoch Kontext.[109]

Es kann festgestellt werden, dass die Vergütungsberichte sowohl in deutscher als auch in englischer Sprache ein komplexes Berichtsinstrument darstellen. Trotz des durch ARUG II erheblich erweiterten Umfangs hat sich die Lesbarkeit zwar statistisch signifikant verschlechtert, allerdings ist für die Adressaten diese Verschlechterung auf dem bestehenden Berichtsniveau praktisch kaum wahrnehmbar.

[109] *Quick et al.* (2024, S. 434) berichten für Risiko- und Chancenberichte einen durchschnittlichen LIX-Wert von 64,90. *Tanirak/Lazar* (2024, S. 17) ermitteln für ESG-Berichte der DAX-40-Unternehmen einen Mittelwert von 59. *Gros/Hanke* (2021, S. 536) finden für IFRS-Konzernabschlüsse Mittelwerte von 63,64 und 63,67. Für KAM-Angaben berichten *Diederich/Velte* (2019, S. 464) einen durchschnittlichen FRE_{Amstad}-Wert von 12,73. *Quick et al.* (2024, S. 434) ermitteln einen Durchschnittswert von 19,40, während *Gros/Hanke* (2021, S. 536) für IFRS-Konzernabschlüsse Werte von 12,59 bzw. 13,27 berichten.

3.6 Limitation

Die Nutzung klassischer Lesbarkeitsmaße, wie die des FOG zur Bewertung der Lesbarkeit von Finanzberichten, steht in der Kritik, da er häufig verwendete Fachbegriffe als komplex einstuft, obwohl diese für die Zielgruppe verständlich sein können.[110] So gibt es bspw. sprachspezifische Besonderheiten, etwa im Deutschen, wo lange zusammengesetzte Wörter das Komplexitätsmaß verzerren können, obwohl sie für Fachleser problemlos verständlich sind.[111] Eine mögliche Alternative bietet der Bog-Index, der die Komplexität von Begriffen auf Basis eines spezifischen Wörterbuchs bewertet und nicht primär die Silbenanzahl als Maßstab heranzieht.[112] Im Kontext dieser Argumentation wird auf den unterschiedlichen Kenntnisstand von privaten und institutionellen Investoren Bezug genommen. Insbesondere bei privaten bzw. weniger versierten Anlegern, die nicht per se mit den finanzwirtschaftlichen Fachbegriffen vertraut sind, ist eine transparente sowie verständliche Kommunikation wichtig.[113] Ferner wurden narrative Teile der Berichte bei der Bereinigung von Abbildungen, Tabellen und ähnlichen Formatierungen teilweise vom Analysegegenstand ausgeschlossen. Eine differenziertere Bereinigung erscheint in diesem Zusammenhang sinnvoll.[114]

3.7 Schlussbetrachtung

Die vorliegende Studie zeigt, dass die Berichterstattung über die Organvergütung börsennotierter Aktiengesellschaften mit der Einführung des ARUG II sowohl in Umfang als auch Detailtiefe deutlich zugenommen hat. Trotz dieses erweiterten Berichtsumfangs ist die Lesbarkeit der Vergütungsberichte nur geringfügig verschlechtert. Grundsätzlich ist festzuhalten, dass der Vergütungsbericht – ungeachtet gesetzlicher Reformen – weiterhin ein komplexes Berichtsinstrument darstellt, das für die Adressaten anspruchsvoll bleibt. Die gesetzlichen Anforderungen nach § 162 AktG, die einen klaren und verständlichen Bericht fordern, werden folglich nur bedingt erfüllt. Unternehmen sollten daher gezielt eine verständlichere Sprache und eine klarere Struktur nutzen, um die Nachvollziehbarkeit für Aktionäre zu verbessern.

Umfangreichere Rechenschaftspflichten ermöglichen den Adressaten sowohl einen intra-Unternehmensvergleich mit den Vorperioden desselben Unternehmens als auch einen inter-

[110] Vgl. *Loughran/McDonald* (2014), S. 1645; *Efretuei/Hussainey* (2023).
[111] Vgl. *Fritz/Töws* (2018), S. 66.
[112] Vgl. *Loughran/McDonald* (2014), S. 1645.
[113] Vgl. *Beier-Middelschulte* (2004), S. 74; *Hooghiemstra et al.* (2017), S. 720 f.
[114] Vgl. bspw. Vorgehen von *Hemmings et al.* (2020, S. 584 ff.), welche die Entfernung derartiger Formatierungen lediglich vornahmen, wenn der narrativer Anteil < 85 % betrug.

Unternehmensvergleich mit branchengleichen Wettbewerbern, um die Angemessenheit der Vergütung zu beurteilen. Es zeigt sich allerdings, dass die Auffindbarkeit und Vergleichbarkeit der Berichte durch uneinheitliche Darstellungsformen eingeschränkt sein können. Die Definition einheitlicher Tabellen, die Festlegung klarer Referenzwerte sowie strukturierte Vorgaben für die Darstellung zentraler Vergütungselemente könnten zur besseren Vergleichbarkeit zwischen Unternehmen beitragen. Darüber hinaus lässt die Mehrheit der Unternehmen ihre Berichte freiwillig über die gesetzlichen Anforderungen hinaus prüfen, was die Verlässlichkeit der Berichterstattung stärkt.

Inwieweit eine klarere und sachlich ausgewogenere Darstellung das Vertrauen der Aktionäre stärkt, sollte Gegenstand weiterer Untersuchungen sein. Studien aus den USA deuten bspw. darauf hin, dass die sprachliche Gestaltung das Aktionärsvotum beeinflusst.[115]

[115] Vgl. *Hemmings et al.* (2020), S. 547.

Literaturverzeichnis

Amstad, T. (1978), Wie verständlich sind unsere Zeitungen?, Dissertation, Universität Zürich, Zürich.

Anderson, J. (1983), Lix and rix: Variations on a little-known readability index, in: Journal of Reading, 26. Jg., S. 490–496.

Anzinger, H. M. (2019), Vorstands- und Aufsichtsratsvergütung: Kompetenzverteilung und Offenlegung nach der zweiten Aktionärsrechterichtlinie, in: Zeitschrift für Unternehmens- und Gesellschaftsrecht, 48. Jg., S. 39–96.

Armeli, N./Misterek, F. (2023), Vorstandsvergütungsstudie 2023. Entwicklung der Vergütung von Vorständen aus DAX- und MDAX-Unternehmen im Geschäftsjahr 2022, Düsseldorf 2023, https://www.imu-boeckler.de/fpdf/HBS-008765/p_mbf_report_2023_80.pdf (abgerufen am 25. 1. 2025).

Aust, V. (2019), Qualität oder Quantität? - Kommunikation in der DAX30- und MDAX-Vergütungsberichterstattung von Vorständen, in: Betriebswirtschaftliche Forschung und Praxis, 71. Jg., S. 468–488.

Balsam, S./Boone, J./Liu, H./Yin, J. (2016), The impact of say-on-pay on executive compensation, in: Journal of Accounting and Public Policy, 35. Jg., S. 162–191.

Bannier, C./Reinschmidt, J. (2022), Vorstandsvergütung in DAX- und MDAX-Unternehmen. Wie nachhaltig sind die Vergütungssysteme der börsennotierten deutschen Unternehmen?, in: WPg, 75. Jg., S. 639–645.

Beck, D./Friedl, G./Schäfer, P. (2020), Executive compensation in Germany, in: Journal of Business Economics, 90. Jg., S. 787–824.

Beck Bil-Komm/Grottel (2024), HGB § 285. In: Grottel, B./Justenhoven, P./Kliem, B./Schubert, W. J. (Hrsg.), Beck'scher Bilanz-Kommentar. Handels- und Steuerbilanz : §§ 238-339, 342-342r HGB, 14. Aufl., München 2024.

BeckOGK/Bayer/Scholz (2024), AktG § 162. In: Stilz, E./Veil, R. (Hrsg.), beck-online.Grosskommentar AktG 2024.

Beier-Middelschulte, A. (2004), Finanzkommunikation junger Emittenten, Wiesbaden 2004.

Björnsson, C. H. (1968), Läsbarhet, Stockholm 1968.

Björnsson, C. H. (1983), Readability of Newspapers in 11 Languages, in: Reading Research Quarterly, 18. Jg., S. 480–497.

Bochkay, K./Brown, S. V./Leone, A. J./Tucker, J. W. (2023), Textual Analysis in Accounting: What's Next?*, in: Contemporary Accounting Research, 40. Jg., S. 765–805.

Bond, S. (2025), Gunning Fog Index, http://gunning-fog-index.com/fog.cgi (abgerufen am 25. 1. 2025).

Bonsall, S. B./Leone, A. J./Miller, B. P./Rennekamp, K. (2017), A plain English measure of financial reporting readability, in: Journal of Accounting and Economics, 63. Jg., S. 329–357.

Brennan, N. M./Guillamon-Saorin, E./Pierce, A. (2009), Methodological Insights, in: Accounting, Auditing & Accountability Journal, 22. Jg., S. 789–832.

Bundesministerium der Justiz und für Verbraucherschutz (2019), Gesetz zur Umsetzung der zweiten Aktionärsrechterichtlinie (ARUG II), https://www.bgbl.de/xaver/bgbl/start.xav?startbk=Bundesanzeiger_BGBl&jumpTo=bgbl119s2637.pdf#/text/bgbl119s2637.pdf?_ts=1748620971769 (abgerufen am 25. 1. 2025).

Burg, T./Plietsch, D./Pott, C./Walter, J. (2016), Die Vorstandsvergütung der DAX30-Unternehmen im Kontext von HGB, DCGK und IFRS, in: KoR: internationale und kapitalmarktorientierte Rechnungslegung, 16. Jg., S. 23–32.

Courtis, J. K. (1995), Readability of annual reports: Western versus Asian evidence, in: Accounting, Auditing & Accountability Journal, 8. Jg., S. 4–17.

Grottel, B./Justenhoven, P./Kliem, B./Schubert, W. J. (Hrsg.) (2024), Beck'scher Bilanz-Kommentar. Handels- und Steuerbilanz : §§ 238-339, 342-342r HGB, 14. Aufl., München 2024.

Deutsche Börse AG (2024), Report: Listed Companies Frankfurt Stock Exchange, https://www.deutsche-boerse-cash-market.com/dbcm-de/instrumente-statistiken/statistiken/gelistete-unternehmen (abgerufen am 25. 1. 2025).

Deutscher Bundestag (2019), Entwurf eines Gesetzes zur Umsetzung der zweiten Aktionärsrechterichtlinie (ARUG II). RegBegr. BT-Drs. 19/9739, https://dserver.bundestag.de/btd/19/097/1909739.pdf (abgerufen am 25. 1. 2025).

Diederich, M./Velte, P. (2019), Key Audit Matters im neuen Bestätigungsvermerk. Eine quantitative Berichtsanalyse für HDAX- und SDAX-Unternehmen im Geschäftsjahr 2017, in: Deutsches Steuerrecht, 57. Jg., S. 461–468.

Efretuei, E./Hussainey, K. (2023), The fog index in accounting research: contributions and challenges, in: Journal of Applied Accounting Research, 24. Jg., S. 318–343.

El-Sayed, D. H./Adel, E./Elmougy, O./Fawzy, N./Hatem, N./Elhakey, F. (2021), The influence of narrative disclosure readability, information ordering and graphical representations on non-professional investors' judgment: evidence from an emerging market, in: Journal of Applied Accounting Research, 22. Jg., S. 138–167.

EU-Kommission (2022), Guidelines on the standardised presentation of the remuneration report under Directive 2007/36/EC, as amended by Directive (EU) 2017/828, as regards the encouragement of long-term shareholder engagement, https://ec.europa.eu/transparency/expert-groups-register/core/api/front/document/88046/download (abgerufen am 25. 1. 2025).

Flesch, R. (1948), A new readability yardstick, in: The Journal of applied psychology, 32. Jg., S. 221–233.

Frankel, R./Jennings, J./Lee, J. (2022), Disclosure Sentiment: Machine Learning vs. Dictionary Methods, in: Management science, 68. Jg., S. 5514–5532.

Fritz, D./Tòws, E. (2018), Text Mining and Reporting Quality in German Banks - A Cooccurrence and Sentiment Analysis, in: Universal Journal of Accounting and Finance, 6. Jg., S. 54–81.

Gao, J./Hartmann, F./Zhang, M./Chen, Y. (2023), The impact of CSR performance and CSR disclosure readability on investors' earnings estimates, in: Accounting & Finance, 63. Jg., S. 1157–1186.

Goette, W./Habersack, M./Kalss, S. (Hrsg.) (2024), Münchener Kommentar zum Aktiengesetz, 6. Aufl., München 2024.

Gosselin, A.-M./Le Maux, J./Smaili, N. (2021), Readability of Accounting Disclosures: A Comprehensive Review and Research Agenda*, in: Accounting Perspectives, 20. Jg., S. 543–581.

Götz, A./Friese, N. (2010), Empirische Analyse der Vorstandsvergütung im DAX und MDAX nach Einführung des Vorstandsvergütungsangemessenheitsgesetzes, in: Corporate Finance, 1. Jg., S. 410–420.

Götz, A./Stahl, M. (2023), Vorstandsvergütung in DAX und MDAX 2022. Analyse der Vergütung und der Vergütungselemente, in: Corporate Finance, 14. Jg., S. 193–196.

Götz, A./Stahl, M./Wiedmann, K. (2022), Vorstandsvergütung in DAX und MDAX 2021. Analyse der Vergütungsberichte nach § 162 AktG, in: Corporate Finance, 13. Jg., S. 215–220.

Grigoleit, H. C. (Hrsg.) (2020), Aktiengesetz. Kommentar, 2. Aufl., München 2020.

Grigoleit/Rachlitz (2020), AktG § 162. In: Grigoleit, H. C. (Hrsg.), Aktiengesetz. Kommentar, 2. Aufl., München 2020.

Gros, M./Hanke, A. (2021), Fehleridentifikation im Enforcement-Verfahren - Was kann die Computerlinguistik leisten? Schwerpunktthema: Prüfungswesen, in: Betriebswirtschaftliche Forschung und Praxis, 73. Jg., S. 519–550.

Grottke, M. (2009), Die strukturale Textanalyse als bilanzanalytisches Instrument zur Auswertung von Lageberichten, in: Die Betriebswirtschaft, 69. Jg.

Guay, W./Samuels, D./Taylor, D. (2016), Guiding through the Fog: Financial statement complexity and voluntary disclosure, in: Journal of Accounting and Economics, 62. Jg., S. 234–269.

Gunning, R. (1952), The technique of clear writing, New York 1952.

Hartmann, M./Weißenberger, B. E. (2024), Information overload research in accounting: a systematic review of the literature, in: Management Review Quarterly, 74. Jg., S. 1619–1667.

Heilmann, M./Höfmann, M./Pott, C. (2020), Berichterstattung von Aufsichtsrat und Vorstand. „Der Ton macht die Musik“: eine textimmanente Analyse, in: WPg, 73. Jg., S. 707–714.

Hemmings, D./Hodgkinson, L./Williams, G. (2020), It's OK to pay well, if you write well: The effects of remuneration disclosure readability, in: Journal of Business Finance & Accounting, 47. Jg., S. 547–586.

Heyd, R./Anselmann, D. (2019), Die Beschreibung der Key Audit Matters in den Bestätigungsvermerken ausgewählter Unternehmen. Analyse anhand eines Scoring-Modells, in: Zeitschrift für Bilanzierung, Rechnungswesen & Controlling, 43. Jg., S. 579–584.

Hohmann, F./Schäfer, D. (2020), Impression Management in der Finanzberichterstattung. Wie Unternehmen die Leser der Berichterstattung durch sprachliche Gestaltung beeinflussen können, in: IRZ : Zeitschrift für internationale Rechnungslegung, 15. Jg., S. 461–467.

Hölters, W./Weber, M. (Hrsg.) (2022), Aktiengesetz. Kommentar, 4. Aufl., München 2022.

Hölters/Weber (2022), AktG § 162. In: Hölters, W./Weber, M. (Hrsg.), Aktiengesetz. Kommentar, 4. Aufl., München 2022.

Hooghiemstra, R./Kuang, Y. F./Qin, B. (2017), Does obfuscating excessive CEO pay work? The influence of remuneration report readability on say-on-pay votes, in: Accounting and Business Research, 47. Jg., S. 695–729.

IASB (2017), Disclosure Initiative—Principles of Disclosure. Discussion Paper DP/2017/1, https://www.ifrs.org/content/dam/ifrs/project/disclosure-initative/disclosure-initiative-principles-of-disclosure/discussion-paper/published-documents/discussion-paper-disclosure-initiative-principles-of-disclosure.pdf (abgerufen am 25. 1. 2025).

IDW (2018), Referentenentwurf eines Gesetzes zur Umsetzung der zweiten Aktionärsrechterichtlinie (ARUG II-RefE), bmj.de/SharedDocs/Downloads/DE/Gesetzgebung/Stellungnahmen/2019/12042018_Stellungnahme_Institut-Wirtschaftspruefer_ARUG-II.pdf?__blob=publicationFile (abgerufen am 25. 1. 2025).

Kang, Y./Cai, Z./Tan, C.-W./Huang, Q./Liu, H. (2020), Natural language processing (NLP) in management research: A literature review, in: Journal of management analytics, 7. Jg., S. 139–172.

Kaya, D./Seebeck, A. (2018), Computergestützte Textanalyse von erweiterten Bestätigungsvermerken, in: WPg, 71. Jg., S. 995–1002.

Kolb, S./Nickel, J. R./Riedl, F./Rupp, A. (2022), Handbuch Aktienrechtliche Organvergütung: Vergütungssystem und Vergütungsberichterstattung nach ARUG II, Stuttgart 2022.

Kuthe, T./Reuters, M. (2022), Der aktienrechtliche Vergütungsbericht nach § 162 AktG. Praxisrelevante Hinweise, in: Betriebs-Berater, 77. Jg., S. 2115–2120.

Laksmana, I./Tietz, W./Yang, Y.-W. (2012), Compensation discussion and analysis (CD&A): Readability and management obfuscation, in: Journal of Accounting and Public Policy, 31. Jg., S. 185–203.

Lenhard, W./Lenhard, A. (2022), Berechnung des Lesbarkeitsindex LIX nach Björnson, https://www.psychometrica.de/lix.html (abgerufen am 25. 1. 2025).

Li, F. (2008), Annual report readability, current earnings, and earnings persistence, in: Journal of Accounting and Economics, 45. Jg., S. 221–247.

Lo, K./Ramos, F./Rogo, R. (2017), Earnings management and annual report readability, in: Journal of Accounting and Economics, 63. Jg., S. 1–25.

Lohmann, C./Danisch, C. (2021), Sprachliche Konsistenz in der Unternehmensberichterstattung: Ein Vergleich der sprachlichen Tonalität von finanziellen und nichtfinanziellen Unternehmensberichten, in: Betriebswirtschaftliche Forschung und Praxis, 73. Jg., S. 179–199.

Lohmann, C./Ohliger, T. (2020), Bankruptcy prediction and the discriminatory power of annual reports: empirical evidence from financially distressed German companies, in: Journal of Business Economics, 90. Jg., S. 137–172.

Loughran, T. I./McDonald, B. (2011), When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks, in: The Journal of Finance, 66. Jg., S. 35–65.

Loughran, T. I./McDonald, B. (2014), Measuring Readability in Financial Disclosures, in: The Journal of Finance, 69. Jg., S. 1643–1671.

Loughran, T. I./McDonald, B. (2016), Textual Analysis in Accounting and Finance: A Survey, in: Journal of Accounting Research, 54. Jg., S. 1187–1230.

Manager Magazin (2022), Klatsche für den Bayer-Vorstand. Vergütungsbericht fällt durch, https://www.manager-magazin.de/unternehmen/pharma/bayer-aktionaere-watschen-vorstand-

auf-hauptversammlung-ab-a-87f75621-aa9e-4f15-bfd4-5c07c5fdfc30 (abgerufen am 25. 1. 2025).

Mauritz, C./Nienhaus, M./Oehler, C. (2023), The role of individual audit partners for narrative disclosures, in: Review of Accounting Studies, 28. Jg., S. 1–44.

Mühlberger, M./Woern, M. (2024), ESG-Ziele in der Vorstandsvergütung. Eine empirische Analyse des EURO STOXX 50, in: KoR: internationale und kapitalmarktorientierte Rechnungslegung, 24. Jg., S. 327–336.

MüKoAktG/Kubis (2024), AktG § 120a. In: Goette, W./Habersack, M./Kalss, S. (Hrsg.), Münchener Kommentar zum Aktiengesetz, 6. Aufl., München 2024.

MüKoAktG/Spindler (2024a), AktG § 162. In: Goette, W./Habersack, M./Kalss, S. (Hrsg.), Münchener Kommentar zum Aktiengesetz, 6. Aufl., München 2024.

MüKoAktG/Spindler (2024b), AktG § 87a. In: Goette, W./Habersack, M./Kalss, S. (Hrsg.), Münchener Kommentar zum Aktiengesetz, 6. Aufl., München 2024.

Online Utility (2025), Tests Document Readability, https://www.online-utility.org/english/readability_test_and_improve.jsp (abgerufen am 25. 1. 2025).

Pöferlein, M. (2021), Sentiment Analysis of German Texts in Finance: Improving and Testing the BPW Dictionary, in: Journal of Banking and Financial Economics, 8. Jg., S. 5–24.

Quick, R./Kordisch, J./Schneider, C. (2024), Zum Einfluss der Abschlussprüfungsqualität auf die Lesbarkeit von Risiko- und Chancenberichten, in: Der Konzern, 22. Jg., S. 431–436.

Quick, R./Pappert, N./Friedrich, C./Carlé, T. (2021), Ähnlichkeiten und "Boilerplates" in der Berichterstattung über Key Audit Matters in den Bestätigungsvermerken der DAX-30-Unternehmen, in: Betriebswirtschaftliche Forschung und Praxis, 73. Jg., S. 461–478.

Quick, R./Pappert, N./Rose, J. (2023), Zur Lesbarkeit von Key Audit Matters. Eine empirische Untersuchung zu den Bestätigungsvermerken der HDAX-Unternehmen, in: WPg, 76. Jg., S. 844–851.

Readability Formulas (2024), Average Readability Consensus Calc, https://readabilityformulas.com/calculator-arlc-formula.php (abgerufen am 25. 1. 2025).

Regierungskomission Deutscher Corporate Governance Kodex (2022), Deutscher Corporate Governance Kodex, https://www.dcgk.de//files/dcgk/usercontent/de/download/kodex/220627_Deutscher_Corporate_Governance_Kodex_2022.pdf (abgerufen am 25. 1. 2025).

Rimmelspacher, D./Kliem, B. (2021), Der Vertikalvergleich im Vergütungsbericht nach § 162 AktG, in: WPg, 74. Jg., S. 1000–1008.

Rimmelspacher, D./Roland, S. (2020), Der Vergütungsbericht nach ARUG II, in: WPg, 73. Jg., S. 201–209.

Rupertus, H./Kaiser, F./Bravidor, M. (2017), Auf hohem Niveau schwierig!?: die Lesbarkeit der Finanzberichterstattung im internationalen Vergleich: eine empirische Analyse für Deutschland, Großbritannien und Frankreich, in: KoR: internationale und kapitalmarktorientierte Rechnungslegung, 17. Jg., S. 79–85.

Schick, A. G./Gordon, L. A./Haka, S. (1990), Information overload: A temporal approach, in: Accounting, organizations and society, 15. Jg., S. 199–220.

Schmidt, M./Pacher, S. (2022), ESG-Leistungskriterien in der Vorstandsvergütung. Eine Marktanalyse auf Basis der DAX-Familie, in: Zeitschrift für Corporate Governance, 17. Jg., S. 112–115.

Schöll, P. (2025), Fleschindex - Berechnen, https://www.fleschindex.de/berechnen/ (abgerufen am 25. 1. 2025).

Schutzgemeinschaft der Kapitalanleger (2024), Geplantes Abstimmungsverhalten der SdK auf der ordentliche Hauptversammlung am 27.08.2024, https://api.sdk.org/m/07345d2b-1028-4fbc-8fee-66189d3f8de7 (abgerufen am 25. 1. 2025).

Stilz, E./Veil, R. (Hrsg.) (2024), beck-online.Grosskommentar AktG 2024.

Stöber, M. (2020), Neuerungen im Aktienrecht durch das ARUG II, in: Deutsches Steuerrecht, 58. Jg., S. 391–398.

Stone, G./Parker, L. D. (2013), Developing the Flesch reading ease formula for the contemporary accounting communications landscape, in: Qualitative research in accounting & management, 10. Jg., S. 31–59.

Tanirak, E./Lazar, C. (2024), Analyse der ESG-Berichtspraxis in DAX 40-Unternehmen. Trends, Bewertungen und Lesbarkeit von 2019 bis 2022, in: ESGZ - Fachzeitschrift für Nachhaltigkeit und Recht, 3. Jg., S. 10–19.

Tönnies, A./Tönnies, M. (2024), Nachhaltigkeit in der Vorstandsvergütung börsennotierter Aktiengesellschaften in Deutschland. Ergebnisse einer Stichprobenanalyse von DAX-40 und S-DAX-Unternehmen für das Geschäftsjahr 2021, in: WPg, 77. Jg., S. 224–232.

Tröger, T. H./Walz, U. (2019), Does Say on Pay Matter? Evidence from Germany, in: European company and financial law review, 16. Jg., S. 381–414.

Velte, P. (2020a), Associations between the financial and industry expertise of audit committee members and key audit matters within related audit reports, in: Journal of Applied Accounting Research, 21. Jg., S. 185–200.

Velte, P. (2020b), "Nachhaltige und langfristige" Vorstandsvergütung nach dem ARUG II. Erste Anmerkungen zur zwingenden Einbeziehung von Sozial- und Umweltaspekten nach § 87 I 2 AktG, in: Neue Zeitschrift für Gesellschaftsrecht, 23. Jg., S. 12–15.

Velte, P. (2021), Effect of overlapping audit and compensation committee memberships on the readability of management compensation reports in the German HDAX, in: Journal of managerial issues, 33. Jg., S. 27–48.

Vetter, E./Tielmann, J. (2022), Zum Standort des Vergütungsberichts nach § 162 AktG, in: Neue Zeitschrift für Gesellschaftsrecht, 25. Jg., S. 387–388.

You, H./Zhang, X. (2009), Financial reporting complexity and investor underreaction to 10-K information, in: Review of Accounting Studies, 14. Jg., S. 559–5.

4. Executive Board Gender Diversity and ESG Performance: Evidence from EU Taxonomy Disclosures

4.1 Publication Details

Authors: Jonas Maiwald, Björn Seeck

Abstract: This study examines whether and how executive board gender diversity is associated with environmental implementation performance under the EU Taxonomy framework. Drawing on stakeholder theory, resource dependence theory, and critical mass theory, we posit that executive gender diversity strengthens firms' capacity to operationalize sustainability requirements through internal investment and operational decisions when the board has reached a critical mass of women. While sustainability reporting has expanded rapidly, it remains unclear which governance characteristics are associated with higher levels of firms' taxonomy alignment. To capture sustainability implementation, we conceptualize EU Taxonomy–based implementation ratios at the overall level as well as separately for capital expenditure, operational expenditure, and turnover. We express aligned activities as a share of eligible activities. This holds sectoral scope constant and shows how far firms realize the alignment potential available to them. Drawing on a balanced panel of STOXX Europe 600 firms over the period 2022 to 2024, we document a positive association between executive gender diversity and overall EU Taxonomy implementation. This association is strongest and most robust for capital expenditure- and operational expenditure-based implementation ratios — domains under direct managerial control — and emerges primarily once a critical mass of women on the executive board is reached. Evidence for the turnover-based implementation ratio is weaker, and several associations attenuate in lagged specifications, warranting caution in drawing causal inferences. Taken together, our results indicate that executive board composition is relevant not only to sustainability reporting but also to firms' realized implementation of EU Taxonomy requirements, thereby informing the ongoing European policy debate on the effective implementation of sustainability regulation.

Keywords EU Taxonomy; ESG Performance; ESG Disclosure; Corporate Sustainable Governance; Women on Board

Publication Status: Working Paper.

4.2 Introduction

Sustainability information has become a core input for capital markets and corporate governance in Europe, with the EU Taxonomy Regulation (EU) 2020/852 translating sustainability objectives into auditable and comparable firm-level key performance indicators (KPIs) under Article 8 within a regulatory framework that continues to evolve through recent omnibus initiatives aimed at narrowing the scope and enhancing practical applicability (European Commission, 2025; European Parliament, 2020). While this standardization improves transparency, whether these metrics translate into substantive changes in firms' sustainability implementation depends on corporate governance, especially board decision-making and oversight. In this context, Directive (EU) 2022/2381 underscores the growing regulatory relevance of gender diversity in corporate governance by mandating gender quotas for executive and non-executive leadership positions in European listed firms (European Parliament, 2022a). The directive seeks to address demographic trends and persistent structural barriers in corporate leadership, with the explicit aim of fostering more resilient decision-making and a more sustainable European economy, effectively linking the role of gender and sustainable corporate behavior. We therefore examine whether executive board gender diversity (BGD) is associated with firms' EU Taxonomy implementation and whether the effects are consistent with critical mass logic. Using a balanced panel of EURO STOXX 600 firms for 2022–2024, we examine the relationship between female representation on executive boards and taxonomy implementation ratios at the aggregate level and separately for turnover, capital expenditure (CapEx), and operating expenditure (OpEx).

The roots of sustainability reporting can be traced back to the Brundtland Commission's definition of sustainable development, which emphasized the balance between meeting present needs and safeguarding the opportunities of future generations (Bebbington et al., 2023; World Commission on Environment and Development, 1987). Building on this, the United Nations 2030 Agenda introduced 17 Sustainable Development Goals (SDGs) and 169 targets, which link environmental, social, and economic objectives with a strong financing dimension, emphasizing the mobilization and alignment of public and private capital as a central means of implementation (Bebbington et al., 2017; Bebbington et al., 2023; United Nations, 2015).

A central prerequisite for the success of these initiatives is the availability of reliable and comparable sustainability-related information for investors to allocate capital effectively and support the transition toward a more sustainable economy. Prior and recent research shows that credible, reliable and high-quality disclosure can reduce information asymmetries and agency conflicts and thereby lower the cost of capital (Healy and Palepu 2001; Leuz and Verrecchia 2000; Leuz and Wysocki 2016; Terry et al. 2023). This underscores the view that disclosure, especially sustainability reporting, is not only legitimate but also a practical tool for financial decision-making and governance (Chrzan and Pott 2024).

The European Union (EU) has operationalized this logic through its sustainable finance agenda. The 2018 Action Plan on Financing Sustainable Growth established reorienting capital flows toward sustainable investment, managing ESG (environmental, social and governance)-related risks, and enhancing transparency as explicit policy objectives (Bebbington et al. 2023; Birindelli et al. 2019; European Commission 2018). The 2019 European Green Deal further reframed sustainability, particularly climate neutrality and resource efficiency, as a central element of the EU's growth strategy (Bebbington et al. 2013; European Commission 2019).

At the same time, the EU acknowledged persistent tendencies in corporate practice toward short-term financial performance and high shareholder payouts, which risk undermining the achievement of Green Deal objectives (Bebbington et al., 2023; EY, 2020; Gray, 2006). In response, the 2020 Sustainable Finance Initiative sought to strengthen the integration of sustainability into corporate governance structures and mechanisms (European Commission, 2021; European Parliament, 2020). However, thus far, the governance dimension of this initiative has focused mainly on directors' remuneration (European Commission, 2021). Broader debates on sustainable corporate governance and the corporate social responsibility (CSR) literature suggest that BGD is a particularly important factor in driving long-term value creation and advancing sustainability performance (Birindelli et al., 2019; Byron & Post, 2016; Campopiano et al., 2023; Nicolò et al., 2022).

Accordingly, the EU's sustainable finance framework, comprising the EU Taxonomy Regulation, the Sustainable Finance Disclosure Regulation (SFDR), and the Corporate Sustainability Reporting Directive (CSRD), combines classification with disclosure

(Chrzan & Pott, 2024; Schütze & Sandbaek, 2025). The EU Taxonomy establishes criteria for determining whether an economic activity is environmentally sustainable: it must contribute substantially to at least one of six environmental objectives, avoid significant harm to the others, and comply with minimum safeguards (European Parliament, 2020; Schütze & Sandbaek, 2025). Article 8 of the EU Taxonomy operationalizes these principles in corporate reporting, requiring firms under the CSRD to disclose the shares of turnover, CapEx, and OpEx, which are taxonomy-eligible and taxonomy-aligned, via standardized definitions and templates (European Parliament, 2020, 2022b). This creates a pathway from global policy goals to auditable, comparable firm-level sustainability data, which can be used by investors, supervisors, and boards to direct financial flows toward measurable transitions.

While financial information remains the primary determinant of capital allocation, sustainability information is increasingly valued and priced by investors (Ferguson et al. 2023). Evidence from the introduction of the EU Taxonomy Regulation shows that investors price firms' taxonomy alignment and reallocate capital toward taxonomy-aligned firms, particularly when investor attention is high (Bassen et al. 2025). Consistent with this development, investments in green assets have grown substantially (PRI 2025). Together, these developments imply that firms' implementation of regulatory sustainability requirements can have economically meaningful consequences by affecting access to capital and financing conditions.

Within this architecture, the standardization of sustainability information enhances its reliability and comparability, but its effectiveness in driving actual behavioral and investment changes depends on corporate governance arrangements (Aguilera et al. 2006). Thus, while disclosure makes sustainability performance measurable, governance, which in turn is related to board structure and diversity, shapes whether and how firms translate these insights into long-term value creation and improved ESG outcomes (Bebbington et al. 2023; Prado-Lorenzo and García-Sánchez 2010).

Research on the role of board composition in shaping corporate sustainability reporting and performance has grown considerably over the past two decades (Ben-Amar et al. 2017; García Martín and Herrero 2020; Lioa et al. 2015; Mallin and Michelon 2011). Evidence from U.S. firms, for example, shows that greater BDG in S&P 500 companies is associated with higher ESG disclosure scores (Tamimi and Sebastianelli 2017). More

broadly, studies highlight that corporate governance mechanisms, especially the structure and composition of boards, play a fundamental role in influencing non-financial disclosure practices and the effectiveness of corporate accountability (Arayssi et al. 2020; Giannarakis et al. 2014; Tamimi and Sebastianelli 2017).

Within this line of research, the presence and proportion of women on boards has been identified as a key determinant of ESG disclosure quality and extent. Empirical work suggests that female directors provide diverse experiences, values and stakeholder-oriented perspectives that strengthen transparency and accountability, thereby enhancing ESG disclosure levels and extending sustainability disclosure practices (Arayssi et al., 2020; Ben-Amar et al., 2017; Nicolò et al., 2022; Qureshi et al., 2020; Wasiuzzaman & Wan Mohammad, 2020). Meta-analytic evidence further substantiates a generally positive relationship between women's board representation and corporate social performance, with stronger effects observable in contexts where boards are more gender balanced and where institutional settings foster inclusive governance (Byron & Post, 2016).

Focusing on the banking sector, Birindelli et al. (2019) demonstrate that gender diversity in leadership positions can significantly improve environmental performance, with evidence of a non-linear relationship: only a critical mass of female directors, particularly in combination with female CEOs, strengthens sustainable outcomes. Similarly, recent reviews emphasize that women directors contribute to corporate sustainability strategies and CSR initiatives by bringing unique resources, ethical orientations and stakeholder-focused perspectives to board decision-making, thereby fostering corporate social performance and long-term value creation (Campopiano et al., 2023).

In the European context, Nicolò et al. (2022) provide large-sample evidence that BGD positively influences ESG disclosure, both in the overall and individual ESG dimensions. Their findings suggest that gender-balanced boards are a crucial mechanism for enhancing sustainable corporate governance by broadening decision-making horizons, facilitating access to resources, and aligning corporate strategies with long-term stakeholder interests.

Despite these important insights, the topics of BDG and ESG performance remain underinvestigated. Most studies focus on the links between BGD, CSR performance and

financial performance, whereas systematic analyses of BGDs effect on ESG performance, especially in European contexts, are scarce (Abdelkader et al. 2024; Amorelli and García-Sánchez 2020; Arayssi et al. 2018; Bannò et al. 2023; Nicolò et al. 2022; Sahu et al. 2025).[128] Moreover, prior research reports mixed and at times contradictory findings, ranging from positive to negative or insignificant effects, suggesting that important contextual factors and potential critical-mass dynamics warrant closer scrutiny. To address this concern, we revisit this research question via novel EU Taxonomy data, which are specifically designed to enhance the comparability and reliability of sustainability disclosures. More precisely, we examine the relationship between female representation on corporate boards and firms' ESG performance, as captured by EU Taxonomy–disclosed KPIs, drawing on a balanced panel of 360 firm-year observations from EURO STOXX 600 constituents over the period 2022–2024.

This paper makes three main contributions. First, we introduce EU Taxonomy Article 8 implementation ratios as a novel, auditable, and comparable firm-level proxy for environmental ESG implementation and distinguish between the aggregate ratio and KPI-specific ratios for turnover, CapEx, and OpEx. By doing so, we move beyond conventional ESG ratings and disclosure scores and leverage a regulatory metric that is directly tied to the EU's sustainable finance architecture.

Second, we provide the first large-sample evidence on how executive BGD relates to EU Taxonomy implementation for EURO STOXX 600 firms from 2022–2024. Using a balanced panel of 120 firms (360 firm-year observations), we document that achieving a critical mass of female executives (≥30%) is associated with higher implementation levels for the aggregated ratio as well as for CapEx and OpEx. In contrast, the association with turnover alignment is not robust and disappears once appropriate models for bounded outcome variables and formal tests of the critical mass assumption are applied.

[128] The terms CSR, sustainability, and ESG are often used interchangeably in both the academic literature and practice (e.g. Velte 2025). The choice of terminology frequently depends on the respective research domain, disciplinary perspective, and measurement approach. In this study, we deliberately adopt the ESG terminology, as our analysis is explicitly based on quantitative, regulatorily defined KPIs stemming from EU Taxonomy disclosures. These standardized metrics allow us to conceptualize an implementation ratio that captures firms' actual sustainability implementation rather than broad or self-defined CSR or sustainability activities (see also section 4.4.2).

Third, we contribute to the critical mass literature by combining a prespecified 30% cutoff with a flexible piecewise specification that allows for a separate slope below the threshold, a discontinuity at the cutoff, and a separate slope above it. Our Wald tests support a parsimonious critical mass formulation for the aggregated, CapEx, and OpEx outcomes while rejecting a threshold-type interpretation for turnover. This differentiated pattern helps reconcile mixed findings in prior research and shows that "critical mass" is outcome-specific even within the same regulatory system.

4.3 Theoretical Foundation and Hypotheses

4.3.1 Theoretical Framework

Stakeholder theory suggests that firms are embedded in a wider social context in which legitimacy and long-term value creation depend on addressing the expectations of multiple stakeholders (De Villiers and Dimes 2021; Freeman 1984; Michelon and Parbonetti 2012). BGD is often seen as enhancing this perspective, as women directors tend to be more attuned to ethical, environmental, and social concerns (Campopiano et al. 2023; Zahid et al. 2020). A growing body of empirical work demonstrates that female representation on boards is associated with both higher CSR and ESG disclosure and performance (Carter et al. 2003; Nadeem et al. 2017; Nicolò et al. 2022 Tamimi and Sebastianelli 2017; Valls Martínez et al. 2019). Moreover, the findings remain mixed, as some studies report no or moderate significant effects of female board representation on CSR performance or disclosure (Coffey and Wang 1998; Rao et al. 2012; Rodriguez-Dominguez et al. 2009), whereas others highlight that such effects are contingent on institutional and contextual factors (Byron and Post 2016; Fernandez et al. 2019).

Evidence from European two-tier systems reinforces this view and is of particular relevance to our setting. In Germany and Austria, the principal representatives of the two-tier model, the presence of women on management boards has been linked to improved ESG outcomes, thereby aligning corporate practices more closely with stakeholder demands for sustainability (Velte 2016). Relatedly, gender-diverse boards have been found to improve the quality of materiality disclosure in integrated reporting, enhancing transparency and credibility from stakeholders' perspectives (Velte 2022). These findings indicate that gender diversity not only reshapes board composition but also strengthens the responsiveness of companies to external legitimacy pressures.

Relatedly, in the banking sector, Birindelli et al. (2019) show that female leaders, whether as board members or CEOs, contribute to stronger environmental performance, although the magnitude of this effect appears to depend on the interaction between board compensation and executive leadership. Stakeholder theory thus provides a useful lens for explaining why more diverse boards are particularly responsive to societal pressures for sustainable and transparent practices.

The contribution of women directors is also visible in the quality of disclosure. Research has shown that women directors help reduce information asymmetries between firms and stakeholders, thereby improving ESG transparency (Giannarakis et al. 2014; Healy and Palepu 2001). Moreover, firms with stronger environmental performance tend to provide more extensive carbon disclosures, which is consistent with resource dependence theory (RDT), which views disclosure as a strategic tool for maintaining legitimacy and securing access to capital under heightened stakeholder scrutiny. (Giannarakis et al. 2016).

RDT complements this perspective by emphasizing the role of boards as providers of critical resources such as legitimacy, expertise, and external linkages (De Villiers and Dimes 2021; Filatotchev 2007; Hillman and Cannella 2007; Pfeffer and Salancik 1978). Women directors broaden these linkages by introducing new perspectives, networks, and knowledge, especially in relation to sustainability and ethical governance (Cruz et al. 2019; Fernandez et al. 2019).

Building on Pfeffer's early work (1972, 1973, 1987), boards can be understood as formal mechanisms through which organizations manage interdependencies with critical external actors. The composition of boards thus reflects strategic efforts to secure access to resources and legitimacy rather than being random or purely symbolic (Pfeffer 1973, 1987). In this view, female directors may serve as boundary spanners while linking the firm to socially salient stakeholder groups, thereby reducing environmental uncertainty and enhancing legitimacy.

Moreover, meta-analytical evidence indicates that gender diverse boards strengthen both monitoring and strategy involvement, two core board functions, with particularly strong effects in institutional environments characterized by greater gender parity and strong shareholder protections (Byron and Post 2016). These findings underscore the idea that female directors bring not only symbolic legitimacy but also substantive governance

resources. As Hillman & Dalziel (2003) argue, boards perform dual roles. They monitor management and provide critical resources. Board capital, comprising directors' human and social capital, determines the board's ability to fulfill these functions. Accordingly, female directors enhance both monitoring effectiveness and access to external legitimacy, thereby linking diversity directly to improved ESG governance capacity.

Further reviews of RDT reinforce Pfeffer's idea that boards act as adaptive mechanisms to manage environmental interdependence through five main strategies (mergers, alliances, board appointments, political action, and executive succession) (Hillman et al. 2009). Within this framework, gender diversity can be seen as an internal governance strategy that enhances a firm's capacity to respond to social and regulatory dependencies, particularly those arising from sustainability expectations.

Taken together, stakeholder theory and RDT offer complementary explanations: gender diverse boards tend to be both more responsive to external pressures and better equipped with pivotal resources to address governance challenges.

Critical mass theory adds an important contingency to these perspectives. Kanter's (1977) foundational work shows that when women remain tokens in male-dominated boards, they are subject to heightened visibility, polarization, and assimilation pressures that limit their influence. This dynamic helps explain why studies often report mixed or inconsistent results on the effects of BGD (Bannò et al. 2023; Velte 2015).

Empirical evidence suggests that meaningful influence emerges only once a critical threshold of female representation is reached. Studies show that positive ESG outcomes materialize when the number of women directors exceeds token levels (Amorelli and García-Sánchez 2020; Bannò et al. 2023; De Masi et al. 2021; Manita et al. 2017; Velte 2025). The presence of female CEOs further amplifies these effects, underscoring that structural power and representation work side by side (Birindelli et al. 2019). Conversely, symbolic appointments without substantive influence may undermine effectiveness and create legitimacy struggles (Tremblay 2016).

Recent work has also highlighted contextual factors. The benefits of female representation appear strongest in the form of favorable organizational conditions, such as strong market positions, where boards are better able to leverage legitimacy and stakeholder expectations (Yang et al. 2025). At the same time, research emphasizes that

sustainability cannot be integrated effectively if female directors are included only symbolically. Embedding sustainability in boardroom practices requires women's voices to be normalized and structurally empowered (Yuriev and Boiral 2024).

In this sense, critical mass theory refines stakeholder and RDT perspectives on sustainable corporate governance by underscoring that the positive influence of women directors on ESG outcomes depends not only on their presence but also on sufficient representation and institutional support.

4.3.2 Literature and Hypothesis Development

Research on the link between board composition and sustainability outcomes and disclosure is well established in the literature (e.g., Arayssi et al. 2020; Giannarakis et al. 2014; Liao et al. 2015; Michelon and Parbonetti 2012; Nicoló et al. 2022; Tamimi and Sebastianelli 2017). Drawing on stakeholder theory, boards orchestrate the firm's response to heterogeneous stakeholder claims and set disclosure policies that sustain organizational legitimacy and capital access over time (Campopiano et al. 2023; Daily et al. 2003; Roberts 1992; Snider et al. 2003; Tamimi and Sebastianelli 2017; Zhang et al. 2013). Hence, board attributes should be reflected in the quality and extent of sustainability information and related performance outcomes (Michelon and Parbonetti 2012). In this view, sustainability disclosure is part of a dialog that reduces information asymmetry and aligns expectations, and it is shaped by who sits on the board and the resources they bring.

RDT complements this perspective, as directors provide access to the knowledge, networks and legitimacy required to navigate evolving sustainability demands (Daily et al. 2003). A more diverse board, specifically with greater female representation, broadens these linkages and tends to be more responsive to social and environmental issues, improving monitoring and strategic involvement in ways that translate into more robust sustainability practices and disclosures (Ben-Amar et al. 2017; Campopiano et al. 2023; Liao et al. 2015; Nicolò et al. 2022). Empirical evidence in European settings shows that a greater proportion of female directors is associated with stronger non-financial and ESG transparency, which is consistent with boards leveraging diversity to address stakeholders' long-term interests (Nicolò et al. 2022; Tamimi and Sebastianelli 2017).

Building on this theoretical foundation, the EU sustainable finance architecture has translated broad policy goals into auditable firm-level KPIs. Article 8 of the EU Taxonomy requires undertakings to disclose the shares of turnover, CapEx and OpEx, which are taxonomy-eligible and taxonomy-aligned, thereby standardizing definitions, formats and templates. This standardization aims explicitly at improving transparency and comparability, curbing greenwashing, and supporting the redirection of capital toward sustainable activities (Chrzan & Pott, 2024; Schütze & Sandbaek, 2025). In capital markets, such quantified, uniform KPIs are intended to sharpen investor assessment and induce firms to improve their taxonomy-aligned performance (Chrzan & Pott, 2024).

These regulatory objectives make the EU Taxonomy implementation ratios (aligned/eligible) a meaningful proxy of (environmental) ESG performance in our setting. It is important to distinguish between the Taxonomy as a reporting instrument and the measure we derive from it. The Taxonomy itself standardizes disclosure rather than performance, and the absolute share of aligned activities largely reflects a firm's business model composition. Our implementation ratios address this concern by construction. By expressing aligned activities as a proportion of eligible activities, the measure conditions on the set of activities that could in principle be aligned and thereby isolates the degree to which a firm converts eligible activities into aligned ones. Firms operating in sectors with limited eligibility are consequently not penalized, and the ratio captures implementation intensity rather than sectoral positioning. On this basis, we treat the implementation ratios as a proxy for environmentally oriented implementation performance, while acknowledging that they remain a regulatory construct whose relationship to physical environmental outcomes is indirect.

Importantly, the prior literature links female board representation to both disclosure intensity and environmental outcomes. Studies using data retrieved from the carbon disclosure project (CDP) and broader ESG metrics find that the likelihood and extensiveness of climate-related (GHG) reporting increase with the percentage of women on boards and that BGD is associated with stronger environmental performance indicators (including fewer environmental lawsuits), which is consistent with enhanced monitoring and stakeholder orientation (Ben-Amar et al. 2017; Birindelli et al. 2019; Liao et al. 2015). While some studies emphasize critical mass dynamics, suggesting that women directors can exert substantive influence only once their presence exceeds a certain

threshold, the overall empirical evidence remains predominantly positive. Recent reviews show that once boards reach a balanced or tilted composition, as Kanter (1977) names it, typically with approximately three women or 30–40% female representation, the effect of gender diversity on CSR and ESG outcomes becomes both significant and durable (Ben-Amar et al. 2017; Birindelli et al. 2019; Velte 2025). This pattern is particularly pronounced in European governance contexts, where regulatory frameworks and stakeholder expectations strongly emphasize diversity, sustainability, and board accountability (Nicolò et al. 2022).

Our empirical models focus on the relative share of women on the board as the main predictor (Buallay et al. 2022; Caby et al. 2024; Enciso-Alfaro and García-Sánchez 2024; García-Sánchez et al. 2024; Velte 2025); consistent with our design, we account for financial controls (size, profitability, leverage) and governance characteristics (e.g., committee structures, independence) that the prior literature has associated with disclosure and ESG outcomes (García-Martín and Herrero 2020; Liao et al. 2015; Nicolò et al. 2022; Tamimi and Sebastianelli 2017).

Within the EU-Taxonomy reporting system, firms disclose eligibility and alignment at the KPI level (turnover, CapEx, OpEx). These categories carry distinct strategic implications, as alignment in turnover reflects current environmentally sustainable business models; alignment in CapEx captures forward-looking investment in environmentally sustainable activities; and alignment in OpEx indicates operational expenditure consistent with environmentally sustainable objectives. Because executive boards determine investment priorities and resource allocation, a more gender-diverse board can be associated with higher implementation ratios across these KPIs, particularly for CapEx, where strategic investment choices and long-term value creation are most visible in the taxonomy framework (Chrzan and Pott 2024; Schütze and Sandbaek 2025).

A conceptual clarification is warranted before we state our hypotheses. The literature reviewed above examines predominantly supervisory or unitary boards and emphasizes monitoring, independence, and legitimacy as the primary transmission channels. Our focus on executive boards departs from this tradition deliberately. In two-tier systems, the supervisory board oversees and appoints, whereas the management board allocates resources, approves investment decisions, and determines the operational scope of regulatory compliance. Since taxonomy alignment is realized through capital and

operating expenditure decisions rather than through oversight, the executive level constitutes the theoretically appropriate locus of analysis. A growing body of work supports this shift in focus. Burkhardt et al. (2020) explicitly extend board-level findings to top management and report higher environmental performance in firms with greater female representation at the executive level, while Saeed et al. (2022) document a positive association between top management gender diversity and the adoption of environmental standards. Caby et al. (2024) adopt a multiteam system perspective and find that reaching a 30% threshold within top management teams improves carbon disclosure quality, indicating that critical mass dynamics operate at the executive level as well. Studies examining both levels simultaneously report complementary rather than substitutive roles (Haque et al. 2026), although the relative strength of each level remains contested, with some evidence suggesting a stronger board effect (Ouni et al. 2020) and other work pointing to distinct ESG dimensions being addressed at each level (Sandretto et al. 2025). We therefore treat executive gender diversity as a construct in its own right rather than as a proxy for supervisory board composition, while acknowledging that the mechanisms established for supervisory boards may operate differently at the executive level.

Taken together, theory and prior findings support a directional prediction that higher relative female representation on the executive board is positively associated with taxonomy implementation, both in the aggregate and for the specific KPIs. We therefore state the following hypotheses:

H_1 Executive BGD is positively associated with the overall EU Taxonomy implementation ratio.

H_2 Executive BGD is positively associated with the EU Taxonomy implementation ratio for turnover.

H_3 Executive BGD is positively associated with the EU Taxonomy implementation ratio for capital expenditures.

H_4 Executive BGD is positively associated with the EU Taxonomy implementation ratio for operating expenditures.

The relative share of women on the board of directors is expected to be positively related to firms' EU-Taxonomy implementation ratios, both in aggregate and across the individual KPIs (turnover, CapEx, and OpEx). This expectation rests on the view that gender-diverse boards enhance decision quality, ethical orientation, and stakeholder-responsive strategies, thereby strengthening firms' ability and willingness to align activities with the EU's sustainability criteria (Liao et al. 2015; Michelon and Parbonetti 2012; Nicolò et al. 2022; Schütze and Sandbaek 2025).

In particular, Taxonomy alignment in CapEx captures forward-looking investment decisions, where diverse boards are more likely to promote long-term sustainable strategies (Ben-Amar et al. 2017; Birindelli et al. 2019). Alignment in turnover reflects the extent to which firms have already integrated sustainable activities into their business models; here, gender-diverse boards may drive credible, stakeholder-oriented reporting and strategic repositioning (Chrzan and Pott 2024; Tamimi and Sebastianelli 2017). Finally, alignment in OpEx indicates operational expenditures consistent with environmental safeguards, which are reinforced by boards' ethical orientation and accountability structures typically associated with greater female representation (Michelon and Parbonetti 2012).

Importantly, these hypotheses are not without tension. While prior studies point to a positive association between board gender diversity and sustainability outcomes, several countervailing mechanisms suggest that such effects may be weak, conditional, or even absent in the context of EU Taxonomy implementation. First, the highly technical and compliance-driven nature of taxonomy reporting may limit the scope for board-level influence, shifting effective control to specialized sustainability, legal, or finance functions within the firm. In this setting, directors' characteristics may matter less than organizational capabilities and regulatory expertise. Second, firms with strong EU Taxonomy alignment may already be subject to intense regulatory and market pressure, making implementation largely a function of external constraints rather than internal governance attributes. Taken together, these considerations suggest that the relationship between executive board gender diversity and EU Taxonomy implementation is theoretically ambiguous, underscoring the need for systematic empirical analysis.

4.4 Research Design

4.4.1 Sample Selection

All the data were retrieved from the LSEG database and included firms listed in the EURO STOXX 600 index as of 15.09.2025. LSEG provides EU Taxonomy information in two formats: "as reported" and "estimated" values generated by LSEG's proprietary algorithms to fill disclosure gaps. As part of the data validation process, we conducted a targeted cross-check of both formats against original company publications. These checks revealed occasional discrepancies between LSEG's estimated EU Taxonomy KPIs and the values reported in companies' official disclosures. Because our study requires verifiable regulatory data and aims to avoid algorithmically imputed EU Taxonomy metrics, we rely exclusively on "as reported" information. Firms for which only estimated EU Taxonomy values were available were therefore excluded from the sample. Firms with missing ESG and governance data (240 firms) were subsequently excluded. In addition, Swiss and UK companies were removed from the sample, as these companies are not subject to the EU Taxonomy regulation due to the absence of EU membership. Financial sector firms, such as banks and insurance companies, were also excluded since they are subject to distinct disclosure requirements, including the mandatory reporting of the Green Asset Ratio from 2024 onward. After applying these exclusions, the final sample comprises 120 firms, resulting in 360 firm-year observations for the period 2022--2024. The final sample is balanced, as all the firms are represented across all three years. The sample firms originate from the following 13 countries: Austria, Belgium, Denmark, Finland, France, Germany, Italy, Luxembourg, the Netherlands, Poland, Portugal, Spain, and Sweden.

Table 4.1 presents the industry composition of the sample. The 120 firms are distributed across 23 industries, indicating substantial sectoral breadth. The largest group belongs to Industrial Goods, followed by Energy/Fossil Fuels and Utilities.

Industries	Frequencies (by ID) Absolute	Frequencies (by Obs.) Absolute	Frequencies Relative (%)
Applied Resources	15	5	4.17
Automobiles & Auto Parts	18	6	5.00
Chemicals	9	3	2.50
Cyclical Consumer Products	12	4	3.33
Cyclical Consumer Services	6	2	1.67
Energy/Fossil Fuels	36	12	10.00
Food & Beverages	3	1	8.83
Food & Drug Retailing	9	3	2.50
Healthcare Services & Equipment	15	5	4.17
Industrial & Commercial Services	27	9	7.50
Industrial Goods	69	23	19.17
Investment Holding Companies	3	1	0.83
Mineral Resources	21	7	5.83
Personal & Household Products & Services	3	1	0.83
Pharmaceuticals & Medical Research	12	4	3.33
Real Estate	15	5	4.17
Renewable Energy	3	1	0.83
Retailers	9	3	2.50
Software & IT Services	15	5	4.17
Technology Equipment	15	5	4.17
Telecommunications Services	6	2	1.67

Transportation	9	3	2.50
Utilities	30	10	8.33
All industries	360	120	100

Table 4.1: Sample composition by industry

4.4.2 Variable Description

4.4.2.1 Dependent Variable

Following the presentation of our sample and the data preparation undertaken to ensure suitability for the research objectives, we distinguish four dependent variables aligned with the hypotheses developed. As no prior study has linked EU Taxonomy disclosures to ESG or environmental performance or examined the individual KPIs of EU Taxonomy disclosure in detail, all four variables were newly operationalized within the scope of this study. The variable aggregated implementation ratio captures the aggregate measure across all three KPIs, expressing the ratio of aligned to eligible activities overall. For the remaining dependent variables, we calculate the ratio of aligned values to eligible values separately for turnover (turnover implementation ratio), CapEx (CapEx implementation ratio), and OpEx (OpEx implementation ratio). Each calculated ratio can range from 0 to 1. This conditional construction is deliberate. Absolute alignment shares would primarily reflect the extent to which a firm operates in Taxonomy-relevant sectors, whereas the ratio of aligned to eligible activities holds sectoral scope constant and captures the extent to which a firm realizes the alignment potential available to it. The measure is therefore comparable across firms with heterogeneous business models, which is a prerequisite in a sample spanning 23 industries.

4.4.2.2 Independent and Control Variables

Consistent with prior research on sustainable corporate governance and non-financial reporting (e.g., Arayssi et al. 2020; Campopiano et al. 2023; Nicolò et al. 2022; Tamimi and Sebastianelli 2017), the key independent variable in this study is executive BGD, captured through a critical mass dummy variable (Exec_Gender_30) that equals 1 when

the share of female executive board members reaches or exceeds 30% and 0 otherwise.[129] This operationalization includes a critical mass perspective and reflects the view that a minimum representation threshold is necessary for women directors to exert substantive influence on board deliberations and sustainability-related decision-making (De Masi et al. 2021; Liu et al. 2014; Lucas-Pérez et al. 2015; Nadeem 2020; Simpson et al. 2010). Gender-diverse executive boards are assumed to improve stakeholder responsiveness and strategic monitoring, thereby fostering greater implementation of EU Taxonomy through greater alignment with environmentally sustainable activities.

To mitigate omitted variable bias, we include several control variables capturing firm specific, governance-related, and contextual characteristics known to influence sustainability performance.

In terms of financial characteristics, we include profitability (ROE), which represents the ratio of net income to total equity, accounting for the firm's financial capacity to allocate resources toward sustainability investments. Leverage (LEV_TA), defined as total debt divided by total assets, captures the degree of financial constraint and creditor monitoring, which may either enhance disclosure incentives or limit environmental investments depending on debt levels (Jizi 2017). Firm size (Size) is measured as the natural logarithm of total assets and controls for visibility and resource availability effects. Larger firms are generally more exposed to stakeholder and regulatory scrutiny and thus more likely to engage in proactive sustainability practices (Roberts 1992; Tamimi and Sebastianelli 2017; Valls Martínez et al. 2019).

The second group captures corporate governance mechanisms. CSR_Committee is a dummy variable equal to 1 if a dedicated sustainability or ESG committee exists at the board or executive level, indicating the board's environmental responsiveness and its inclination for sustainability matters (Birindelli et al. 2019; Cucari 2017; Liao et al. 2015). CSR_Audit is a dummy variable equal to 1 when the firm's sustainability or non-financial report is externally assured, indicating enhanced reliability and transparency of sustainability processes (Mio et al. 2020). Board independency (Board_Independency) is measured as the share of independent directors on the supervisory board. Independent

[129] In all cases we used two-way clustered standard errors (industry×firm or country×firm) and, for fractional outcomes, Papke–Wooldridge fractional logit; results were qualitatively unchanged, supporting the jump interpretation at 30%.

directors are expected to strengthen monitoring quality and accountability toward non-financial stakeholders (Jizi 2017; Zahid et al. 2020). Furthermore, Sustainability_Comp is a dummy variable that equals 1 if executive remuneration includes ESG- or sustainability-linked performance metrics. This dummy reflects the integration of sustainability objectives into managerial incentive schemes and signals alignment with long-term sustainability goals (Berrone and Gomez-Mejia 2009; Birindelli et al. 2019; Hong et al. 2016; Tamimi and Sebastianelli 2017).

Finally, we control for industry and time effects by including sector dummies (factor(industry)) and year fixed effects (factor(year)). Industry controls capture heterogeneity in environmental exposure and regulatory intensity across sectors, whereas year effects absorb common shocks related to macroeconomic and regulatory developments, most notably the progressive implementation of the EU Taxonomy from 2023–2024.

Together, these variables allow the model to isolate the independent effect of executive BGD on firms' degree of EU Taxonomy implementation while controlling for financial structure, governance quality, and temporal or industry-specific heterogeneity. A concise overview of all the independent and control variables, including their operational definitions and expected signs, is provided in *Table 4.7* in the Appendix.

4.4.3 Model Specification

4.4.3.1 Baseline Specification and Model Choice

A Hausman test (Onali et al. 2017) fails to reject orthogonality between regressors and firm effects; hence, random effects are consistent and preferred on efficiency grounds. We therefore report random effects as the main estimator. A Breusch–Pagan Lagrange multiplier test rejects Var (u_i) =0, indicating non-egligible firm heterogeneity; pooled OLS is therefore inappropriate.[130]

With a short time dimension (T = 3), we explicitly diagnose error dependence along both panel dimensions. We implement the Papke-Wooldridge test for first-order serial correlation within firms (Papke and Wooldridge 1996) and the Pesaran CD test for cross-

[130] For completeness, we also estimated pooled OLS models. While the coefficients for executive gender diversity were positive and statistically significant, the OLS estimates are sensitive to functional-form assumptions. For the econometric reasons outlined above, we therefore rely on the random-effects and fractional-logit models as the more appropriate specifications.

sectional dependence across firms (Pesaran 2015, 2021). Consistent with the presence of common shocks and within-unit persistence, our baseline model includes year and industry fixed effects to absorb time-specific aggregate shocks and time-invariant industry heterogeneity (Caskey and Ozel 2019). Furthermore, our model reports two-way cluster-robust standard errors at the firm and industry levels (Cameron et al. 2011; Petersen 2009). To improve small-sample inference given the finite number of clusters, we employ CR2 variance estimators (Imbens and Kolesár 2016; Pustejovsky and Tipton 2018) with Satterthwaite degrees of freedom (Imbens and Kolesár 2016; Satterthwaite 1946).[131]

In the following, we present the baseline random-effects panel regression model.

$$y_{j,it} = \alpha_j + \beta_j ExecGender_{i,t} + \sum_{k=1}^{K} \gamma_{jk}\, Controls_{k,it} + \eta_{s(i)} + \lambda_t + u_i + \varepsilon_{j,it} \quad (1)$$

where

- $y_{j,it}$ denotes the EU-Taxonomy implementation ratio of firm *i* in year *t* for outcome *j*, with $j \in \{1,2,3,4\}$ aggregated, Turnover, CapEx, and OpEx implementation ratios, respectively;
- $Controls_{k,it}$ includes ROE, LEV_TA, Size, CSR_Committee, CSR_Audit and Board_Independency;
- $\eta_{s(i)}$ are industry fixed effects;
- λ_t represents year fixed effects;
- u_i is a firm-specific random intercept;
- $\varepsilon_{i,t}$ is the idiosyncratic error.

We calculate variance inflation factors (VIFs) for single-parameter regressors and adjusted generalized VIFs for multiparameter factors. All the underlying VIF and adjusted GVIF values are below 2 and therefore far below the commonly cited critical

[131] As an additional robustness check, we also estimated all models with one-way clustered standard errors by industry, by firm, and by year. The coefficients and significance levels remain qualitatively unchanged, and in some cases even become stronger, indicating that our conclusions are not driven by the specific clustering choice.

threshold of 10, indicating that multicollinearity is not a concern in our specifications (O´Brian, 2007).

4.4.3.2 Critical Mass Modeling

Motivated by the literature on critical thresholds, we model nonlinearities via a prespecified 30% cutoff (Kanter 1977).[132] Our critical mass model (CMM) sets the effect below the threshold to zero and does not allow for discontinuity at the cutoff; only above the threshold does the effect increase with the share:

$$y_{j,it} = \alpha_j + \beta_j x_{>c,it} + \sum_{k=1}^{K} \gamma_{jk}\, Controls_{k,it} + \eta_{s(i)} + \lambda_t + u_i + \varepsilon_{j,it} \tag{2}$$

where

$$x_{>c,it} \equiv ExecGender_{it} \cdot 1[ExecGender_{it} \geq c],\ c = 0{,}30 \tag{3}$$

To validate these restrictions, we first estimate a greater number of slopes below and above the threshold and a potential discontinuity at the cutoff with the flexible piecewise specification:

$$y_{j,it} = \alpha + \beta_{0j} x_{<c,it} + \beta_{1j} D_{c,it} + \beta_{2j} x_{>c,it} + \sum_{k=1}^{K} \gamma_{jk}\, Controls_{k,it} + \eta_{s(i)} + \lambda_t + u_i + \varepsilon_{j,it} \tag{4}$$

where

$$x_{<c,it} = ExecGender_{it} \cdot 1[ExecGender_{it} < c] \tag{5}$$

$$D_{c,it} = 1\ [ExecGender_{it} \geq c] \tag{6}$$

$$x_{>c,it} = ExecGender_{it} \cdot 1[ExecGender_{it} \geq c] \tag{7}$$

We then test the joint null hypothesis $H_0: \beta_0 = \beta_1 = 0$ via a cluster-robust Wald test based on the same two-way VCOV. Since H_0 cannot be rejected, CMM represents a parsimonious, data-supported specification.

[132] Sensitivity analyses are conducted using 25% and 35% as alternative thresholds.

4.5 Results

4.5.1 Descriptive Statistics

The descriptive statistics in *Table 4.2* and *Table 4.3* summarize the characteristics of the 120 firms observed over the 2022–2024 period (360 firm-year observations). The implementation ratios show considerable variation, with CapEx exhibiting the highest average implementation (0.294) and OpEx the lowest (0.198). Female representation on the executive board averages 21.1%, with substantial cross-sectional dispersion (SD 0.144) and values ranging from 0% to 81.8%. Board independence is comparatively high (mean 60.4%). With respect to governance structures, approximately 90% of firms maintain a CSR committee, and a similar proportion have their sustainability reporting externally assured.

Table 4.2: Descriptive Statistics for Continuous Variables

Variables	Obs.	Mean	Std. Dev	Min	Max
Impl_Aggregated (DV)	360	0.267	0.335	0	1
Impl_Turnover (DV)	360	0.272	0.367	0	1
Impl_CapEx (DV)	360	0.294	0.375	0	1
Impl_OpEx (DV)	360	0.198	0.336	0	1
Exec_Gender	360	0.211	0.144	0	0.818
ROE	360	14.05	12.79	-40.531	70.456
LEV_TA	360	26.843	14.591	0.683	72.821
Size	360	23.866	1.416	20.814	27.295
Board_Independency[133]	360	60.351	22,058	0	100
Number of ID	120	120	120	120	120

[133] The minimum value of 0 for Board Independency is driven by five observations from two firms that are recorded as having no independent directors in the Refinitiv dataset. A robustness check excluding these two firms yields only marginal changes in the estimated coefficients. Board Independency becomes weakly significant ($p < 0.10$) in Model (4) for OpEx, while all other results remain qualitatively unchanged.

Table 4.3: Descriptive Statistics of the Independent Variables (Dummies)

	Value	2022	2023	2024	Total	%
CSR_Committee	0	16	13	6	35	9.72
	1	104	107	114	325	90.28
	Total	120	120	120	360	100
CSR_Audit	0	14	12	9	35	9.72
	1	106	108	111	325	90.28
	Total	120	120	120	360	100
Sustainability_Comp	0	34	24	24	82	22.78
	1	86	96	96	278	77.22
	Total	120	120	120	360	100

Observations (n) and shares (%), n = 360 firm-year observations

4.5.2 Random Effects Model Results

Table 4.4 presents the random-effects estimations, including industry and year fixed effects and two-way clustered (industry × firm) standard errors. For the turnover-based implementation ratio, the flexible piecewise specification shows that neither the slope below 30% nor a discrete change at the 30% threshold is significantly different from zero. The Wald test therefore indicates that the data do not support a threshold-type pattern for turnover. When the predictor is split into segments (< 30% and ≥ 30%), the coefficients become imprecise, and the initially significant baseline association disappears. This means that the positive effect observed in the baseline model does not reflect a structural change at 30% but rather a smooth average association over the full range of female representation. Consequently, the critical mass formulation is not applicable for turnover, and it is appropriate to report the baseline model alongside the segmented version. In contrast, for CapEx, OpEx, and the aggregated implementation ratio, the critical mass specification is the relevant model for interpretation.

Across all model variants, the coefficient of Exec Gender (30), indicating that firms with at least 30 percent female executives, is positive and statistically significant. Statistically

significant associations are consistently observed for the OpEx-based implementation ratio ($\beta = 0.199$, $p = 0.0179$), the aggregated implementation ratio ($\beta = 0.206$, $p = 0.0695$) and the CapEx-based implementation ratio ($\beta = 0.253$, $p = 0.042$) in the critical mass specification. For the turnover-based implementation ratio, the coefficient remains positive and statistically significant in the baseline model ($\beta = 0.350$, $p = 0.0454$).

These findings support the expected positive association for the overall implementation measure (H_1) as well as for the turnover (H_2), CapEx (H_3), and OpEx-based (H_4) implementation measures. Importantly, while the turnover association is positive in the baseline model, additional tests indicate that this effect should not be interpreted as evidence of a discontinuous critical mass shift at 30% but rather as an average positive association.

Among the governance controls, CSR_Committee shows a negative coefficient that becomes significant only for OpEx ($\beta = -0.102$, $p < 0.0855$), whereas CSR_Audit, Board Independency, and Sustainability Compensation remain insignificant. Year dummies are consistently positive—especially for 2024—reflecting the general rise in environmental alignment following EU-Taxonomy adoption.

Table 4.4: Random Effects Model

	CMM (H_1) Aggregated	Baseline (H_1) Aggregated	CMM (H_2) Turnover	Baseline (H_2) Turnover	CMM (H_3) CapEx	Baseline (H_3) CapEx	CMM (H_4) OpEx	Baseline (H_4) OpEx
Exec_Gender (30)	0.206*	n/a	0.242*	n/a	0.253**	n/a	0.199**	n/a
	(0.104)		(0.117)		(0.112)		(0.0727)	
Exec_Gender	n/a	0.156	n/a	0.350**	n/a	0.157	n/a	0.267*
		(0.128)		(0.157)		(0.141)		(0.144)
ROE	-0.0005	-0.0005	-0.0004	-0.0002	-0.001	-0.001	0.0002	0.0003
	(0.0011)	(0.0012)	(0.0008)	(0.0009)	(0.0014)	(0.0014)	(0.0012)	(0.0012)
LEV_TA	-0.0003	0	-0.0012	-0.001	-0.0008	-0.0004	-0.0021	-0.0019
	(0.0022)	(0.0021)	(0.0021)	(0.0021)	(0.0026)	(0.0024)	(0.0013)	(0.0013)
Size	-0.0143	-0.0171	-0.0261	-0.0311	-0.0123	-0.0153	0.0007	-0.0031
	(0.0278)	(0.0266)	(0.0309)	(0.0317)	(0.0323)	(0.0302)	(0.023)	(0.022)
CSR_Committee	-0.0755	-0.0749	-0.051	-0.0488	-0.0289	-0.0285	-0.102*	-0.102
	(0.0583)			(0.0344)				

		(0.0564)	(0.0346)		(0.0874)	(0.0847)	(0.0493)	(0.0474)
CSR_Audit	-0.0392	-0.043	-0.0346	-0.0471	-0.0106	-0.0134	-0.0367	-0.0487
	(0.0488)	(0.0494)	(0.0319)	(0.0358)	(0.0482)	(0.0487)	(0.0672)	(0.0673)
Board_Independency	-0.0013	-0.0014	-0.0007	-0.0009	-0.0019	-0.0019*	0.0001	0
	(0.001)	(0.001)	(0.0012)	(0.0013)	(0.0011)	(0.001)	(0.0009)	(0.0009)
Sustainability_COMP	-0.0333	-0.0027	-0.0703	-0.066	-0.0158	-0.0157	-0.0568	-0.0556
	(0.03)	(0.0289)	(0.0597)	(0.0581)	(0.0456)	(0.0446)	(0.0377)	(0.0386)
FE Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE Industry	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	360	360	360	360	360	360	360	360
R^2	0.2996	0.2927	0.2701	0.2705	0.2913	0.2819	0.2914	0.2927

Notes: Standard errors are in parentheses. ***, **, and * denote significance at the 1, 5 and 10% levels, respectively. Standard errors are two-way clustered at the firm and industry levels. Model fit is satisfactory for a short balanced panel ($R^2 \approx 0.27–0.29$) with moderate cross-sectional heterogeneity.

4.6 Additional Analyses and Robustness Checks

4.6.1 Additional Analysis: Interaction with CO_2-Intensive Companies

To examine whether the association between executive board gender diversity and EU Taxonomy implementation depends on firms' environmental profiles, we incorporate firm-level CO_2 intensity into the analysis. CO_2 intensity is measured as Scope 1 and 2 emissions relative to revenue and averaged at the firm level over the sample period.[134] On this basis, we construct a high-intensity dummy High_CO_2 (mean) that equals one for firms whose average CO_2 intensity exceeds the sample mean and zero otherwise. We then augment the baseline specifications by including High_CO_2 (mean) and its interaction with executive gender diversity, thereby testing whether the effect of board gender diversity on EU Taxonomy implementation is systematically stronger or weaker in more carbon-intensive firms.

$$y_{j,it} = \alpha_j + \beta_j ExecGender_{i,t} + \delta_j CO2Int_{it} + (ExecGender_{it} x\ CO2Int_{it}) + \sum_{k=1}^{K} \gamma_{jk}\, Controls_{k,it} + \eta_{s(i)} + \lambda_t + u_i + \varepsilon_{j,it} \quad (8)$$

Table 4.5 provides the results. High_CO_2 (mean) is negatively associated with EU Taxonomy implementation in several specifications, indicating that more carbon-intensive firms tend to exhibit a lower EU Taxonomy implementation ratio. However, the interaction terms between executive gender diversity (including the 30% critical mass indicator) and High_CO_2 (mean) are consistently small and statistically insignificant in both the random effects and fractional logit models (see section 4.6.2). This finding suggests that gender-diverse executive boards are associated with greater Taxonomy implementation irrespective of firms' CO_2 intensity levels, whereas carbon intensity primarily shifts the overall level of alignment rather than moderates the diversity effect. As a robustness check, we replicate the analysis via a quartile-based high-intensity dummy High_CO_2 (Q4) for firms in the top CO_2-intensity quartile; the results remain qualitatively unchanged and do not reveal any economically meaningful interaction effects.

[134] We rely on Scope 1 and Scope 2 emissions due to data availability and measurement concerns associated with collecting Scope 3 emissions (Blanco 2016; Mahieux et al. 2025; Patchell 2018; WRI and WBCSD 2011). This choice is also consistent with the EU Taxonomy's emphasis on firms' own activities and operational control.

Table 4.5: Random Effects Model: Main Results vs. CO2 Intensity

	CMM (H_1) Aggregated	Baseline (H_1) Aggregated	CMM (H_2) Turnover	Baseline (H_2) Turnover	CMM (H_3) CapEx	Baseline (H_3) CapEx	CMM (H_4) OpEx	Baseline (H_4) OpEx
(MAIN) Exec_Gender (30)	0.206* p=0.0695 (0.104)	n/a	0.242* p=0.0604 (0.117)	n/a	0.253** p=0.042 (0.112)	n/a	0.199** p=0.0179 (0.0727)	n/a
(MAIN) Exec_Gender	n/a	0.156 p=0.247 (0.128)		0.350** p=0.0454 (0.157)	n/a	0.157 p=0.288 (0.141)	n/a	0.267* p=0.0898 (0.144)
Exec_Gender (30)	0.149* p= 0.0592 (0.0699)	n/a	0.183* p= 0.0859 (0.0963)	n/a	0.164* p= 0.055 (0.0756)	n/a	0.184 p= 0.0661 (0.0889)	n/a
Exec_Gender	n/a	0.111 p=0.353	n/a	0.291* p=0.0734	n/a	0.065 p=0.540	n/a	0.213 p= 0.246

		(0.111)		(0.146)		(0.103)		(0.173)
	-0.196	-0.214	-0.316**	-0.351**	-0.174	-0.238	-0.171	-0.210
High_CO_2 (mean)	p=0.171	p=0.300	p= 0.0367	p= 0.0451	p= 0.138	p=0.225	p= 0.406	p= 0.366
	(0.117)	(0.185)	(0.102)	(0.132)	(0.0939)	(0.172)	(0.184)	(0.211)
	0.196	0.144	0.226	0.259	0.357	0.409	-0.015	0.178
Exec_Gender x HighCO_2	p= 0.647	p=0.799	p= 0.398	p= 0.477	p= 0.418	p= 0.529	p= 0.924	p= 0.553
	(0.407)	(0.530)	(0.249)	(0.334)	(0.411)	(0.599)	(0.150)	(0.279)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE Industry	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	360	360	360	360	360	360	360	360
R^2	0.3196	0.3153	0.3075	0.3077	0:3024	0.2970	0.3052	0.3083

Notes: Standard errors are in parentheses. ***, **, and * denote significance at the 1, 5 and 10% levels, respectively. Standard errors are two-way clustered at the firm and industry levels

4.6.2 Robustness Check: Fractional Logit Model

To address potential biases arising from the bounded nature of the dependent variables, we reestimate all the models via a fractional logit specification (Papke and Wooldridge 1996). The implementation measures are defined as percentages ranging from 0--100, rescaled to the [0,1] interval, with a considerable mass at both bounds. Specifically, the share of zero outcomes ranges from 31% (Total) to 61% (OpEx), whereas between 3% and 9% of observations take the upper boundary value of one. In such settings, linear models may produce predictions outside the feasible range and heteroskedastic residuals, potentially biasing standard errors even under robust inference. Fractional-logit quasi-maximum likelihood estimation (QMLE) provides consistent estimates for proportion data, properly accounting for the asymmetric and bounded distributions while preserving interpretability in terms of average marginal effects (Gallani et al. 2015; Wooldridge 2010).[135]

Table 4.6 reports the results of this specification, employing the same covariates, industry and year fixed effects, and two-way (industry × firm) clustered standard errors as in the linear model. For the aggregate implementation measure as well as for the CapEx and OpEx indicators, the positive and significant effect of Exec Gender (30) remains stable in sign and significance, indicating that the critical mass relationship is not an artifact of linearity assumptions or heteroskedasticity. These results demonstrate that the implementation effects associated with achieving a 30% share of female executives are robust to alternative functional forms and to the explicit modeling of the bounded nature of the dependent variables.

Given the high concentration of observations at the boundaries (0 and 1) and the sensitivity of the QMLE approach to such mass points, the absence of significance in the fractional logit model suggests that the turnover effect observed in the baseline model is likely driven by functional-form assumptions rather than a stable underlying pattern. Moreover, the fractional logit specification reveals a few minor differences in control significance. Leverage and sustainability compensation become significant in the OpEx model, and the aggregate (Total) specification gains slightly stronger significance for

[135] Accordingly, numerous studies in accounting employ this method, including Caskey and Ozel (2019), Anantharaman et al. (2018), Chen. et al. (2015), and Li (2013).

Exec Gender (30). These shifts are consistent with the higher sensitivity of the QMLE estimator to boundary observations (0/1) and confirm that the main effect of gender diversity persists even when extreme outcome values receive greater weight.

The robustness analysis yields a differentiated assessment of the hypotheses. The evidence that executive board gender diversity is positively associated with the overall EU Taxonomy implementation ratio (H_1) is strongly supported and remains intact under the fractional logit specification. Similarly, the positive associations for the CapEx-based implementation ratio (H_3) and the OpEx-based implementation ratio (H_4) are robust across linear and nonlinear models. In contrast, the evidence for the turnover-based implementation ratio (H_2) is sensitive to model specification: while a linear association emerges in the random effects model, the fractional logit results indicate that effects are concentrated at higher levels of gender diversity, which is consistent with a critical mass pattern. Given this sensitivity, conclusions regarding turnover-based implementation should be drawn with caution.

As a further robustness check, we reestimate our preferred random effects and fractional logit specifications using a one-year lag of executive board gender diversity instead of the contemporaneous share. This lag structure reduces concerns about reverse causality and aligns board composition with subsequent EU Taxonomy implementation outcomes while retaining industry and year fixed effects and two-way clustered standard errors.

Overall, the lagged estimates remain positive in sign and are broadly consistent with the main results, although their statistical strength is slightly weaker. The association is most robust for OpEx, which remains positive and statistically significant across estimators (with the critical mass variant somewhat less precisely estimated). For CapEx, the coefficients remain positive but are generally weaker, with only marginal support under the critical mass specification. In contrast, effects for the aggregate implementation ratio become statistically insignificant, and turnover shows at most sporadic marginal significance.

Table 4.6: Fractional Logit Model

	CMM (H_1) Aggregated	Baseline (H_1) Aggregated	CMM (H_2) Turnover	Baseline (H_2) Turnover	CMM (H_3) CapEx	Baseline (H_3) CapEx	CMM (H_4) OpEx	Baseline (H_4) OpEx
Exec_Gender (30)	1.89**	n/a	2.78*	n/a	2.08**	n/a	2.41**	n/a
	(0.764)		(1.21)		(0.829)		(0.696)	
Exec_Gender	n/a	1.79		3.36	n/a	1.70	n/a	2.96*
		(1.42)		(1.90)		(1.50)		(1.37)
ROE	-0.0067	0.0072	-0.0071	-0.0086	-0.0106	-0.0111	0.0002	-0.0006
	(0.012)	(0.0119)	(0.0156)	(0.0155)	(0.0128)	(0.0125)	(0.017)	(0.0166)
LEV_TA	-0.0114	0.0094	-0.0095	-0.0074	-0.0142	-0.0119	-0.0206**	-0.017*
	(0.0153)	(0.0139)	(0.0144)	(0.0133)	(0.0199)	(0.0179)	(0.007)	(0.0079)
Size	-0.0831	-0.112	-0.131	-0.172	-0.0578	-0.0864	-0.0099	-0.0478
	(0.184)	(0.167)	(0.202)	(0.193)	(0.219)	(0.195)	(0.166)	(0.148)
CSR_Committee	-0.503	-0.482	-0.514	-0.492	-0.170	-0.143	-0.514	-0.477
	(0.640)	(0.579)	(0.614)	(0.581)	(0.904)	(0.836)	(0.323)	(0.229)

CSR_Audit	-0.156	-0.239	-0.209	-0.425	0.159	0.0921	-0.338	-0.489
	(0.361)	(0.375)	(0.370)	(0.382)	(0.419)	(0.418)	(0.804)	(0.816)
Board_Independency	-0.0021	-0.0024	0.0083	0.0078	-0.0084	-0.0084	0.0102	0.0093
	(0.0049)	(0.0051)	(0.0089)	(0.0089)	(0.0063)	(0.0066)	(0.0097)	(0.0091)
Sustainability_COMP	-0.308	-0.292	-0.909	-0.9088	-0.297	-0.269	-0.976**	-1.01**
	(0.335)	(0.34)	(0.620)	(0.661)	(0.389)	(0.379)	(0.383)	(0.372)
FE Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE Industry	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	360	360	360	360	360	360	360	360
Pseudo-R^2 (Efron)	0.5025	0.4952	0.5222	0.5128	0.4804	0.4688	0.4810	0.4882

Notes: Standard errors in parentheses. ***, **, and * denote significance at the 1, 5 and 10% levels, respectively. Standard errors are two-way clustered at the firm and industry levels.

4.7 Discussion

Grounded in stakeholder theory, RDT, and critical mass theory, our theoretical framework posited that more gender-diverse executive boards would be associated with higher levels of sustainability implementation by improving stakeholder responsiveness, broadening access to sustainability-relevant resources, and fostering inclusive decision-making processes. Consistent with these expectations, our results show that firms with at least 30% female executives exhibit significantly higher levels of EU Taxonomy implementation in the aggregate KPI as well as for CapEx and OpEx. While CapEx and OpEx exhibit a clear and robust critical mass dynamic, the turnover indicator shows no stable association across model specifications.

Stakeholder theory emphasizes that legitimacy and long-term value creation depend on firms' ability to respond to heterogeneous stakeholder expectations (Freeman, 1984; Michelon & Parbonetti, 2012). Gender-diverse boards are theorized to enhance this responsiveness by integrating broader ethical and environmental considerations into corporate decision-making (Campopiano et al., 2023; Zahid et al., 2020). Our findings extend this perspective beyond conventional disclosure and reputation effects. Strong CapEx and OpEx results under the critical mass specification suggest that the association between executive gender diversity and implementation is concentrated in internal domains, that is, in areas where stakeholder expectations are operationalized through resource allocation, capital budgeting, and operational adjustments. . In this sense, the EU Taxonomy implementation ratios capture not only boards' willingness to respond to stakeholder demands but also their capacity to translate these demands into concrete investment and process decisions. These results indicate, in line with further research, that diversity becomes materially effective only once women reach a representation level that allows meaningful participation in strategic and resource-related decisions.

From an RDT perspective, boards function as providers of critical resources, legitimacy, and linkages to the external environment (De Villiers and Dimes 2021; Hillman and Cannella 2007; Hillman and Dalziel 2003; Pfeffer and Salancik 1978). Female directors expand this resource base by contributing diverse cognitive frames, access to sustainability networks, and legitimacy among social and institutional stakeholders (Bannò et al. 2023; Campopiano et al. 2023; Cruz et al. 2019; Jizi 2017). Our evidence aligns with this logic: the observed performance improvements in CapEx and OpEx implementation are consistent with the notion that gender-diverse executive boards mobilize and integrate such resources more effectively. They appear better able to channel financial capital and organizational capital into sustainability-oriented initiatives.

Critical mass theory provides an additional lens for interpreting these patterns. Building on Kanter's (1977) distinction between token and balanced groups, our findings corroborate the proposition that meaningful influence emerges only once women surpass a threshold of approximately 30%. Above the threshold, gender diversity becomes substantively rather than symbolically effective, enabling boards to embed sustainability considerations into investment portfolios and operational processes (Amorelli and García-Sánchez 2020; Bannò et al. 2023; Byron and Post 2016; De Masi et al. 2021; Kanter 1977; Velte 2025).

The turnover dimension requires a different and distinct interpretation. Unlike CapEx and OpEx, which reflect decisions under direct managerial control, turnover is tied to existing business models, product portfolios, and market demand, which change only gradually and are strongly shaped by the current activity mix underlying EU Taxonomy eligibility and alignment (Chrzan and Pott 2024; Schütze and Sandbaek 2025). Given this structural inertia, a direct, contemporaneous effect of gender diversity on turnover-based implementation is theoretically less plausible. Consistent with this view, our empirical results indicate that while a linear association emerges in baseline random-effects models, the relationship becomes sensitive to model specification. More flexible nonlinear specifications suggest that turnover-based implementation may respond primarily at higher levels of gender diversity, in line with a critical mass pattern. Accordingly, evidence for turnover-related implementation is weaker and should be interpreted with caution. This suggests that influence on turnover-related implementation (if present) likely materializes only with a time lag, as firms realign their activity portfolios and generate taxonomy-conforming turnover streams through longer-term strategic shifts.

It is therefore reasonable that the association is strongest for internal implementation domains such as OpEx and CapEx — where management has immediate control over investment and process decisions — rather than on market-driven outcomes that depend on factors outside managerial discretion. This interpretation is consistent with evidence showing that governance quality and top-management composition shape firms' strategic orientation and sustainability-related decision-making, which in turn affects how environmental and investment priorities are integrated into corporate practices (Ben-Amar et al. 2017; Birindelli et al. 2019; Byron and Post 2016; Hillman and Dalziel 2003; Post et al. 2011; Rao and Tilt 2016).

Our additional analyses further sharpen this interpretation. First, CO_2 intensity is negatively associated with EU Taxonomy implementation in several specifications, but the interaction between CO_2 intensity and executive gender diversity is consistently insignificant. This finding indicates that gender-diverse

executive boards are associated with greater implementation irrespective of whether firms operate in carbon-intensive settings. CO_2 intensity is related to the overall level of EU Taxonomy implementation but does not systematically alter the relationship between BGD and implementation.

Second, in lagged specifications, the association is most stable for OpEx, whereas effects for the aggregated ratio and CapEx weaken and turnover remains inconsistent. This pattern is compatible with a sequencing logic, according to which diversity-related governance changes might first affect operational spending practices, then investment planning, and only later market-based outcomes. Given the three-year observation window, however, it is equally compatible with the alternative that the contemporaneous association reflects a common underlying orientation that attenuates once the predictor is temporally ordered. We return to this ambiguity in the limitations section.

A third observation concerns two governance controls whose coefficients run counter to conventional expectations. The presence of a CSR committee is negatively signed throughout and reaches significance for OpEx, while sustainability-linked executive compensation is negatively and significantly associated with OpEx implementation in the fractional logit specification. Both variables are typically expected to signal sustainability commitment and should therefore correlate positively with implementation. Our findings suggest otherwise, at least for operational expenditure. One interpretation is that these instruments function primarily as legitimacy devices rather than steering mechanisms. This reading is consistent with evidence that ESG criteria in remuneration schemes are frequently non-binding, carry limited weight in overall pay, and lend themselves to symbolic adoption (Maas 2016; Haque & Ntim, 2020; Bebchuk & Tallarita, 2022; Lu 2023). Similarly, the near-universal presence of CSR committees in our sample, with approximately 90% of firms maintaining one, limits its discriminatory power and may indicate institutional isomorphism rather than substantive governance. We caution against over-interpreting these coefficients given the limited sample and the concentration of observations at the distribution boundaries. Nevertheless, the pattern warrants attention, as it suggests that formally institutionalized sustainability governance does not by itself translate into implementation, whereas the composition of the executive board does.

Overall, the results provide empirical support for stakeholder theory and RDT while introducing execution capacity as an additional transmission mechanism linking gender diversity to sustainability performance.

Previous research has focused mainly on symbolic and communicative channels and the notion that gender-diverse boards enhance ESG disclosure, transparency, and reputation (e.g., Arayssi et al. 2020;

Giannarakis et al. 2014; Mallin and Michelon 2011; Michelon and Parbonetti 2012; Nicolò et al. 2022; Tamimi and Sebastianelli 2017). Our findings complement this literature by shedding light on the less examined implementation side of sustainability performance and reporting. In the context of the EU's strengthened sustainable-finance framework, which increasingly relies on market-based mechanisms to drive the green transition, we contribute to a more nuanced understanding of how executive gender diversity may relate not only to reporting practices but also to firms' capacity to meet regulatory sustainability requirements in practice. The results indicate that diversity relates not only to what firms report but also to how far they progress in aligning internal activities with EU Taxonomy criteria.

This execution channel complements rather than replaces legitimacy-based mechanisms. This finding suggests that diversity enhances cognitive variety, decision quality, and access to sustainability-relevant resources, thereby enabling boards to translate stakeholder expectations and regulatory requirements into concrete environmental action. This perspective aligns with RDT's view of boards as adaptive mechanisms managing organizational dependencies and extends stakeholder theory by showing that governance inclusiveness co-occurs with tangible operational outcomes under the EU Taxonomy framework. Taken together, the results indicate that the EU sustainable finance architecture provides a setting in which board gender diversity becomes visible not only in disclosure practices but also in the pattern of environmental implementation across key Taxonomy KPIs.

Several limitations qualify the interpretation of our findings and point to avenues for future research.

The most consequential limitation concerns causal inference. Our random-effects and fractional logit specifications identify associations rather than causal effects, and the direction of the relationship cannot be established within our design. The observed patterns are consistent with three explanations: gender-diverse executive boards may strengthen firms' capacity to implement Taxonomy requirements, as our theoretical framework suggests; firms already advanced in sustainability implementation may subsequently appoint more women to executive positions, whether because such firms hold broader diversity commitments or because they attract candidates with sustainability expertise, among whom women are more strongly represented; or both characteristics may reflect a common underlying organizational orientation. Our lagged specifications offer only partial reassurance: while the OpEx association remains stable, the aggregated ratio loses significance and CapEx weakens substantially once the predictor is temporally ordered — a pattern consistent with reverse causality or common-cause explanations — and a one-year lag is in any case short relative to the pace at which board composition and investment strategies change. We therefore refrain from causal claims and interpret our results as evidence of systematic co-occurrence between executive

gender diversity and Taxonomy implementation. Establishing causality requires designs beyond the reach of our three-year panel: longer time series that reveal whether changes in board composition precede or follow changes in implementation, event studies around executive appointments, or identification strategies exploiting exogenous variation, such as the staggered introduction of national gender quotas across European jurisdictions.

A second limitation concerns measurement. Our implementation ratios are derived from a regulatory reporting framework rather than from physical environmental outcomes. While the conditional construction of the ratio holds sectoral scope constant and thereby mitigates concerns about business model heterogeneity, it cannot establish that higher alignment corresponds to lower emissions or reduced environmental pressure. In addition, the three available reporting years restrict us to a level-based analysis; longer time series would allow implementation to be measured dynamically as the change in aligned shares, capturing firms' implementation trajectories rather than their state at a point in time.

A third limitation relates to the scope of our governance construct. We measure gender diversity at the executive level only, although prior work links both supervisory and executive boards to sustainability outcomes (Haque et al., 2024; Sandretto et al., 2025). Examining both levels jointly, including their interaction, would provide a fuller account of how governance composition relates to implementation and would allow a direct test of whether the mechanisms established for supervisory boards operate similarly at the executive level.

Finally, our sample comprises 120 large listed firms from 13 European countries. While this design ensures comparability and data quality, it limits generalizability to smaller firms, non-listed entities, and non-European regulatory contexts. As Taxonomy reporting matures and the reporting population evolves under the Omnibus framework, larger and more heterogeneous samples will permit broader generalization and more differentiated subsample analyses.

4.8 Conclusion

This study provides the early empirical evidence on how executive BGD is related to firms' environmental implementation performance under the EU Taxonomy framework. By linking established governance theories to a novel regulatory performance metric, this study shows that a more balanced gender composition in top management enhances firms' capacity to operationalize sustainability objectives, particularly through capital and operational expenditures under direct managerial control. Executive gender diversity is most strongly associated with the CapEx- and OpEx-

based and overall implementation ratios, with effects being most pronounced once a critical mass of women is reached, whereas the turnover-based implementation ratio displays weaker and less robust associations, reflecting its dependence on existing business models, portfolio composition, and market demand and thus more gradual and delayed adjustment dynamics. These findings suggest that gender diversity operates primarily through internal governance, resource allocation, and execution mechanisms, whereas market-based implementation outcomes respond more slowly.

Overall, this study contributes to emerging debates on corporate governance and sustainable finance by providing evidence that board composition relates not only to transparency and legitimacy but also for the substantive implementation of sustainability frameworks such as the EU Taxonomy. Against the backdrop of recent Omnibus initiatives aimed at simplifying sustainability reporting and enhancing practical applicability, our findings remain relevant by highlighting the role of internal governance structures in shaping firms' implementation of the EU Taxonomy.

Appendix

Table 4.7: Variable Definitions, Measurements and Descriptions

Variable	Definition/Measurement	Description
Exec_Gender_30	Dummy variable equal to 1 when the share of female executive board members ≥ 30%; 0 otherwise	Captures executive BGD based on the critical mass concept, assuming substantive influence on board deliberations once women reach a minimum representation threshold.
Exec_Gender_CMOnly_25/ Exec_Gender_CMOnly_35	Alternative thresholds of 25% and 35% used for robustness checks	Tests sensitivity of the critical mass specification; contrasted with a flexible piecewise variant
Size	Natural logarithm of total assets	Controls for visibility and resource availability; larger firms face higher stakeholder and regulatory scrutiny.
ROE	Net income divided by total equity	Indicates firms' profitability and financial capacity to invest in sustainability initiatives.
LEV_TA	Total debt divided by total assets	Reflects leverage and financial constraints that may influence sustainability investments and disclosure behavior.

CSR_Committee	Dummy variable equal to 1 if a dedicated sustainability or ESG committee exists at board or executive level	Indicates the board's environmental responsiveness and reflects the formalization of sustainability oversight structures within corporate governance, as such committees enhance engagement with ESG matters.
CSR_Audit	Dummy variable equal to 1 if the sustainability or non-financial report is externally assured	Indicates external verification and enhances credibility of sustainability disclosures.
Board_Independency	Share of independent directors on the supervisory board	Measures the independence of the board and its monitoring strength over management with respect to non-financial performance.
Sustainability_Comp	Dummy variable equals to 1 if executive remuneration includes ESG- or sustainability-linked performance metrics.	Captures the integration of sustainability objectives into managerial incentive schemes and signals alignment with long-term sustainability goals.
factor(industry)	Industry dummies	Control for cross-sector heterogeneity in environmental exposure and regulatory context.

factor(year)	Year fixed effects (2023–2024)	Capture time-specific shocks and EU Taxonomy implementation effects.

References

Abdelkader, M. G., Gao, Y., and Elamer, A. A. 2024. "Board Gender Diversity and ESG Performance: The Mediating Role of Temporal Orientation in South Africa Context." *Journal of Cleaner Production* 440: 140728. https://doi.org/10.1016/j.jclepro.2024.140728.

Aguilera, R. V., C. A. Williams, J. M. Conley, and D. E. Rupp. 2006. "Corporate Governance and Social Responsibility: A Comparative Analysis of the UK and the US." Corporate Governance: An International Review 14(3): 147–58. https://doi.org/10.1111/j.1467-8683.2006.00495.x.

Amorelli, M.-F., and I.-M. García-Sánchez. 2020. "Trends in the Dynamic Evolution of Board Gender Diversity and Corporate Social Responsibility." Corporate Social Responsibility and Environmental Management 28(2): 537–54. https://doi.org/10.1002/csr.2079.

Anantharaman, D., and E. C. Chuk. 2018. "The Economic Consequences of Accounting Standards: Evidence from Risk-Taking in Pension Plans." The Accounting Review 93(4): 23–51. https://doi.org/10.2308/accr-51937.

Arayssi, M., M. Jizi, and H. H. Tabaja. 2020. "The Impact of Board Composition on the Level of ESG Disclosures in GCC Countries." Sustainability Accounting, Management and Policy Journal 11(1): 137–61. https://doi.org/10.1108/SAMPJ-05-2018-0136.

Bannò, M., E. Filippi, and S. Trento. 2023. "Women in Top Echelon Positions and Their Effects on Sustainability: A Review, Synthesis and Future Research Agenda." Journal of Management & Governance 27(1): 181–251. https://doi.org/10.1007/s10997-021-09604-7.

Bassen, A., O. Kordsachia, K. Lopatta, and W. Tan. 2025. "Revenue Alignment with the EU Taxonomy Regulation in Developed Markets." Journal of Banking & Finance 170: 107339. https://doi.org/10.1016/j.jbankfin.2024.107339.

Bebbington, J., S. Russell, and I. Thomson. 2017. "Accounting and Sustainable Development: Reflections and Propositions." Critical Perspectives on Accounting 48: 21–34. https://doi.org/10.1016/j.cpa.2017.06.002.

Bebbington, J., M. Laine, C. Larrinaga, and G. Michelon. 2023. “Environmental Accounting in the European Accounting Review: A Reflection.” European Accounting Review 32(5): 1107–28. https://doi.org/10.1080/09638180.2023.2254351.

Bebchuk, L. A., and R. Tallarita. 2022. "The Perils and Questionable Promise of ESG-Based Compensation." Journal of Corporation Law 48, no. 1: 37–74.

Ben-Amar, W., M. Chang, and P. McIlkenny. 2017. “Board Gender Diversity and Corporate Response to Sustainability Initiatives: Evidence from the Carbon Disclosure Project.” Business Strategy and the Environment 26(6): 832–50. https://doi.org/10.1002/bse.1956.

Berrone, P., and L. R. Gomez-Mejia. 2009. “Environmental Performance and Executive Compensation: An Integrated Agency-Institutional Perspective.” Academy of Management Journal 52(1): 103–26. https://doi.org/10.5465/amj.2009.36461950.

Birindelli, G., A. P. Iannuzzi, and M. Savioli. 2019. “The Impact of Women Leaders on Environmental Performance: Evidence on Gender Diversity in Banks.” Corporate Social Responsibility and Environmental Management 26(6): 1485–99. https://doi.org/10.1002/csr.1762.

Blanco, C., F. Caro, and C. J. Corbett. 2016. “The State of Supply Chain Carbon Footprinting: Analysis of CDP Disclosures by US Firms.” Journal of Cleaner Production 135: 1189–97. https://doi.org/10.1016/j.jclepro.2016.06.132.

Buallay, A., R. Hamdan, E. Barone, and A. Hamdan. 2022. “Increasing Female Participation on Boards: Effects on Sustainability Reporting.” International Journal of Finance & Economics 27(1): 111–24. https://doi.org/10.1002/ijfe.2141.

Burkhardt, K., P. Nguyen, and E. Poincelot. 2020. "Agents of Change: Women in Top Management and Corporate Environmental Performance." Corporate Social Responsibility and Environmental Management 27, no. 4: 1591–1604. https://doi.org/10.1002/csr.1907.

Byron, K., and C. Post. 2016. “Women on Boards of Directors and Corporate Social Performance: A Meta-Analysis.” Corporate Governance: An International Review 24(4): 428–42. https://doi.org/10.1111/corg.12165.

Caby, J., C. Coron, and Y. Ziane. 2024. “How Does Gender Diversity in Top Management Teams Affect Carbon Disclosure and Its Quality? Evidence from the Technological Industry.” Technological Forecasting and Social Change 199: 123077. https://doi.org/10.1016/j.techfore.2023.123077.

Cameron, A. C., J. B. Gelbach, and D. L. Miller. 2011. “Robust Inference with Multiway Clustering.” Journal of Business & Economic Statistics 29(2): 238–49. https://doi.org/10.1198/jbes.2010.07136.

Campopiano, G., G. Patricia, and D. Gimenez-Jimenez. 2023. “Women Directors and Corporate Social Performance: An Integrative Review of the Literature and a Future Research Agenda.” Journal of Business Ethics 182(3): 717–46. https://doi.org/10.1007/s10551-021-04999-7.

Carter, D. A., B. J. Simkins, and W. G. Simpson. 2003. “Corporate Governance, Board Diversity, and Firm Value.” The Financial Review 38(1): 33–53. https://doi.org/10.1111/1540-6288.00034.

Caskey, J., and N. B. Ozel. 2019. “Reporting and Non-Reporting Incentives in Leasing.” The Accounting Review 94(6): 137–64. https://doi.org/10.2308/accr-52367.

Chen, C. X., E. M. Matsumura, J. Y. Shin, and S. Y. C. Wu. 2015. “The Effect of Competition Intensity and Competition Type on the Use of Customer Satisfaction Measures in Executive Annual Bonus Contracts.” The Accounting Review 90(1): 229–63. https://doi.org/10.2308/accr-50870.

Chrzan, S., and C. Pott. 2024. “Limiting Environmental Reporting Flexibility: Investor Judgment Based on the EU Taxonomy.” Review of Quantitative Finance and Accounting 63(2): 1511–48. https://doi.org/10.1007/s11156-024-01297-x.

Cruz, C., R. Justo, M. Larraza-Kintana, and L. Garcés-Galdeano. 2019. “When Do Women Make a Better Table? Examining the Influence of Women Directors on Family Firms’ Corporate Social Performance.” Entrepreneurship Theory and Practice 43(2): 282–301. https://doi.org/10.1177/1042258718796080.

Daily, C. M., D. R. Dalton, and A. A. Cannella. 2003. "Corporate Governance: Decades of Dialog and Data." Academy of Management Review 28(3): 371–82. https://doi.org/10.2307/30040727.

De Masi, S., A. Słomka-Gołębiowska, C. Becagli, and A. Paci. 2021. "Toward Sustainable Corporate Behavior: The Effect of the Critical Mass of Female Directors on Environmental, Social, and Governance Disclosure." Business Strategy and the Environment 30(4): 1497–1514. https://doi.org/10.1002/bse.2721.

De Villiers, C., and R. Dimes. 2021. "Determinants, Mechanisms and Consequences of Corporate Governance Reporting: A Research Framework." Journal of Management and Governance 25(1): 7–26. https://doi.org/10.1007/s10997-020-09530-0.

Enciso-Anfaro, S.-Y., and I.-M. García-Sánchez. 2024. "Do Boards Care about Planetary Boundaries? A Gender Perspective on Circular Economy Disclosures." Business Strategy and the Environment 33(5): 4562–97. https://doi.org/10.1002/bse.3700.

European Commission. 2018. Communication from the Commission to the European Parliament, the European Council, the Council, the European Central Bank, the European Economic and Social Committee and the Committee of the Regions: Action Plan: Financing Sustainable Growth.

European Commission. 2019. Communication from the Commission to the European Parliament, the European Council, the Council, the European Central Bank, the European Economic and Social Committee and the Committee of the Regions: The European Green Deal.

European Commission. 2021. Sustainable Corporate Governance Initiative: Inception Impact Assessment Ares (2020)4034032.

European Commission. 2025. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions – Commission Work Programme 2025 COM(2025) 45 Final.

European Parliament. 2020. Regulation (EU) 2020/852 of the European Parliament and of the Council of 18 June 2020 on the Establishment of a Framework to Facilitate Sustainable Investment, and Amending Regulation (EU) 2019/2088.

European Parliament. 2022a. Directive (EU) 2022/2381 of the European Parliament and of the Council of 23 November 2022 on Improving the Gender Balance among Directors of Listed Companies and Related Measures.

European Parliament. 2022b. Directive (EU) 2022/2464 of the European Parliament and of the Council of 14 December 2022 Amending Regulation (EU) No 537/2014, Directive 2004/109/EC, Directive 2006/43/EC and Directive 2013/34/EU, as Regards Corporate Sustainability Reporting.

EY. 2020. Study on Directors' Duties and Sustainable Corporate Governance.

Ferguson, P. J., J. Hronsky, and M. Pinnuck. 2023. "Who Pays Attention to Sustainability Reports and Why? Evidence from Google Search Activity." Accounting & Finance 63(3): 3519–51. https://doi.org/10.1111/acfi.13045.

Fernandez, W. D., M. F. Burnett, and C. B. Gomez. 2019. "Women in the Boardroom and Corporate Social Performance: Negotiating the Double Bind." Management Decision 57(9): 2201–22. https://doi.org/10.1108/MD-08-2017-0738.

Filatotchev, I. 2007. "Corporate Governance and the Firm's Dynamics: Contingencies and Complementarities." Journal of Management Studies 44(6): 1041–56. https://doi.org/10.1111/j.1467-6486.2007.00716.x.

Freeman, R. E. 1984. Strategic Management: A Stakeholder Approach. Pitman.

Gallani, S., R. Krishnan, and J. M. Wooldridge. 2015. Applications of Fractional Response Model to the Study of Bounded Dependent Variables in Accounting Research. Working Paper No. 16–016. Harvard Business School.

García Martín, C. J., and B. Herrero. 2020. "Do Board Characteristics Affect Environmental Performance? A Study of EU Firms." Corporate Social Responsibility and Environmental Management 27(1): 74–94. https://doi.org/10.1002/csr.1775.

García-Sánchez, I.-M., S. Marín-Hernández, E. Ortiz-Martínez, and D. Aibar-Guzmán. 2024. "Diversity, Equity, and Inclusion Reporting in European Union Companies: The Role of

Female Directors and the European Regulatory Framework." Business Strategy and the Environment 33(7): 7021–40. https://doi.org/10.1002/bse.3854.

Giannarakis, G., G. Konteos, and N. Sariannidis. 2014. "Financial, Governance and Environmental Determinants of Corporate Social Responsible Disclosure." Management Decision 52(10): 1928–51. https://doi.org/10.1108/MD-05-2014-0296.

Giannarakis, G., G. Konteos, N. Sariannidis, and G. Chaitidis. 2016. "The Relation between Voluntary Carbon Disclosure and Environmental Performance: The Case of S&P 500." International Journal of Law and Management 59(6): 784–803. https://doi.org/10.1108/IJLMA-05-2016-0049.

Gray, R. 2006. "Social, Environmental and Sustainability Reporting and Organisational Value Creation: Whose Value? Whose Creation?" Accounting, Auditing & Accountability Journal 19(6): 793–819. https://doi.org/10.1108/09513570610709872.

Haque, F., G. Adjei-Mensah, T. H. H. Nguyen, and C. G. Ntim. 2026. "Does Gender Diversity in Corporate Boards and Executive Management Teams Influence Carbon Performance? Evidence from Europe." Accounting Forum 50, no. 2: 215–244. https://doi.org/10.1080/01559982.2024.2423989.

Haque, F., and C. G. Ntim. 2020. "Executive Compensation, Sustainable Compensation Policy, Carbon Performance and Market Value." British Journal of Management 31, no. 3: 525–546. https://doi.org/10.1111/1467-8551.12395.

Healy, P. M., and K. G. Palepu. 2001. "Information Asymmetry, Corporate Disclosure, and the Capital Markets: A Review of the Empirical Disclosure Literature." Journal of Accounting and Economics 31(1): 405–40. https://doi.org/10.1016/S0165-4101(01)00018-0.

Hillman, A. J., and A. A. Cannella. 2007. "Organizational Predictors of Women on Corporate Boards." Academy of Management Journal 50(4): 941–52. https://doi.org/10.5465/amj.2007.26279222.

Hillman, A. J., and T. Dalziel. 2003. "Boards of Directors and Firm Performance: Integrating Agency and Resource Dependence Perspectives." Academy of Management Review 28(3): 383–96. https://doi.org/10.5465/amr.2003.10196729.

Hong, B., Z. Li, and D. Minor. 2016. “Corporate Governance and Executive Compensation for Corporate Social Responsibility.” Journal of Business Ethics 136(1): 199–213. https://doi.org/10.1007/s10551-015-2962-0.

Imbens, G. W., and M. Kolesár. 2016. “Robust Standard Errors in Small Samples: Some Practical Advice.” Review of Economics and Statistics 98(4): 701–12. https://doi.org/10.1162/REST_a_00552.

Jizi, M. 2017. “The Influence of Board Composition on Sustainable Development Disclosure.” Business Strategy and the Environment 26(5): 569–78. https://doi.org/10.1002/bse.1943.

Kanter, R. M. 1977. “Some Effects of Proportions on Group Life: Skewed Sex Ratios and Responses to Token Women.” American Journal of Sociology 82(5): 965–90. https://doi.org/10.1086/226425.

Leuz, C., and R. E. Verrecchia. 2000. “The Economic Consequences of Increased Disclosure.” Journal of Accounting Research 38(Suppl.): 91–124. https://doi.org/10.2307/2672910.

Leuz, C., and P. D. Wysocki. 2016. “The Economics of Disclosure and Financial Reporting Regulation: Evidence and Suggestions for Future Research.” Journal of Accounting Research 54(2): 525–622. https://doi.org/10.1111/1475-679X.12115.

Li, E. X. 2013. “Revealing Future Prospects without Forecasts: The Case of Accelerating Material Contract Filings.” The Accounting Review 88(5): 1769–1804. https://doi.org/10.2308/accr-50505.

Liao, L., L. Luo, and Q. Tang. 2015. “Gender Diversity, Board Independence, Environmental Committee and Greenhouse Gas Disclosure.” The British Accounting Review 47(4): 409–24. https://doi.org/10.1016/j.bar.2014.01.002.

Liu, Y., Z. Wei, and F. Xie. 2014. “Do Women Directors Improve Firm Performance in China?” Journal of Corporate Finance 28: 169–84. https://doi.org/10.1016/j.jcorpfin.2013.11.016.

Lu, L. 2023. "ESG-Based Remuneration in the Wave of Sustainability." Journal of Corporate Law Studies 23, no. 1: 297–339. https://doi.org/10.1080/14735970.2023.2253888.

Lucas-Pérez, M. E., A. Mínguez-Vera, J. F. Baixauli-Soler, J. F. Martín-Ugedo, and G. Sánchez-Marín. 2015. "Women on the Board and Managers' Pay: Evidence from Spain." Journal of Business Ethics 129(2): 265–80. https://doi.org/10.1007/s10551-014-2148-1.

Mahieux, L., H. Sapra, and G. Zhang. 2025. "Measuring Greenhouse Gas Emissions: What Are the Costs and Benefits?" Journal of Accounting Research 63(3): 1063–1105. https://doi.org/10.1111/1475-679X.12613.

Manita, R., M. G. Bruna, R. Dang, and L. Houanti. 2017. "Board Gender Diversity and ESG Disclosure: Evidence from the USA." Journal of Applied Accounting Research 19(2): 206–24. https://doi.org/10.1108/JAAR-01-2017-0024.

Maas, K., and S. Rosendaal. 2016. "Sustainability Targets in Executive Remuneration: Targets, Time Frame, Country and Sector Specification." Business Strategy and the Environment 25, no. 6: 390–401. https://doi.org/10.1002/bse.1880.

Michelon, G., and A. Parbonetti. 2012. "The Effect of Corporate Governance on Sustainability Disclosure." Journal of Management and Governance 16(3): 477–509. https://doi.org/10.1007/s10997-010-9160-3.

Mio, C., M. Fasan, C. Marcon, and S. Panfilo. 2020. "The Predictive Ability of Legitimacy and Agency Theory after the Implementation of the EU Directive on Non-Financial Information." Corporate Social Responsibility and Environmental Management 27(6): 2465–76. https://doi.org/10.1002/csr.1968.

Nadeem, M. 2020. "Does Board Gender Diversity Influence Voluntary Disclosure of Intellectual Capital in Initial Public Offering Prospectuses? Evidence from China." Corporate Governance: An International Review 28(2): 100–118. https://doi.org/10.1111/corg.1230.

Nadeem, M., R. Zaman, and I. Saleem. 2017. "Boardroom Gender Diversity and Corporate Sustainability Practices: Evidence from Australian Securities Exchange Listed Firms." Journal of Cleaner Production 149: 874–85. https://doi.org/10.1016/j.jclepro.2017.02.141.

Nicolò, G., G. Zampone, G. Sannino, and S. De Iorio. 2022. "Sustainable Corporate Governance and Non-Financial Disclosure in Europe: Does the Gender Diversity Matter?" Journal of Applied Accounting Research 23(1): 227–49. https://doi.org/10.1108/JAAR-04-2021-0100.

O'Brien, R. M. 2007. "A Caution regarding Rules of Thumb for Variance Inflation Factors." Quality & Quantity 41(5): 673–90. https://doi.org/10.1007/s11135-006-9018-6.

Onali, E., G. Ginesti, and C. Vasilakis. 2017. "How Should We Estimate Value-Relevance Models? Insights from European Data." The British Accounting Review 49(5): 460–73. https://doi.org/10.1016/j.bar.2017.05.006.

Ouni, Z., J. Ben Mansour, and S. Arfaoui. 2020. "Board/Executive Gender Diversity and Firm Financial Performance in Canada: The Mediating Role of Environmental, Social, and Governance (ESG) Orientation." Sustainability 12, no. 20: 8386. https://doi.org/10.3390/su12208386.

Papke, L. E., and J. M. Wooldridge. 1996. "Econometric Methods for Fractional Response Variables with an Application to 401(k) Plan Participation Rates." Journal of Applied Econometrics 11(6): 619–32. https://doi.org/10.1002/(SICI)1099-1255(199611)11:6<619::AID-JAE418>3.0.CO;2-1.

Patchell, J. 2018. "Can the Implications of the GHG Protocol's Scope 3 Standard Be Realized?" Journal of Cleaner Production 185: 941–58. https://doi.org/10.1016/j.jclepro.2018.03.003.

Pesaran, M. H. 2014. "Testing Weak Cross-Sectional Dependence in Large Panels." Econometric Reviews 34(6–10): 1089–1117. https://doi.org/10.1080/07474938.2014.956623.

Pesaran, M. H. 2021. "General Diagnostic Tests for Cross-Sectional Dependence in Panels." Empirical Economics 60(1): 13–50. https://doi.org/10.1007/s00181-020-01875-7.

Petersen, M. A. 2008. "Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches." The Review of Financial Studies 22(1): 435–80. https://doi.org/10.1093/rfs/hhn053.

Pfeffer, J. 1972. "Size and Composition of Corporate Boards of Directors: The Organization and Its Environment." Administrative Science Quarterly 17(2): 218–29.

Pfeffer, J. 1973. "Size, Composition, and Function of Hospital Boards of Directors: A Study of Organization–Environment Linkage." Administrative Science Quarterly 18(3): 349–65.

Pfeffer, J. 1987. “A Resource Dependence Perspective on Interorganizational Relations.” In Intercorporate Relations: The Structural Analysis of Business, edited by M. S. Mizruchi and M. Schwartz, 22–55. Cambridge University Press.

Pfeffer, J., and G. R. Salancik. 1978. The External Control of Organizations: A Resource Dependence Perspective. Harper & Row.

Post, C., N. Rahman, and E. Rubow. 2011. “Green Governance: Board of Directors’ Composition and Environmental Corporate Social Responsibility.” Business & Society 50(1): 189–223. https://doi.org/10.1177/0007650310394642.

Prado-Lorenzo, J.-M., and I.-M. García-Sánchez. 2010. “The Role of the Board of Directors in Disseminating Relevant Information on Greenhouse Gases.” Journal of Business Ethics 97(3): 391–424. https://doi.org/10.1007/s10551-010-0515-0.

Principles for Responsible Investment. 2025. Global Responsible Investment Trends: Inside PRI Reporting Data. https://www.unpri.org/download?ac=23004.

Pustejovsky, J. E., and E. Tipton. 2018. “Small-Sample Methods for Cluster-Robust Variance Estimation and Hypothesis Testing in Fixed Effects Models.” Journal of Business & Economic Statistics 36(4): 672–83. https://doi.org/10.1080/07350015.2016.1247004.

Qureshi, M. A., S. Kirkerud, K. Theresa, and T. Ahsan. 2020. “The Impact of Sustainability (Environmental, Social, and Governance) Disclosure and Board Diversity on Firm Value: The Moderating Role of Industry Sensitivity.” Business Strategy and the Environment 29(3): 1199–1214. https://doi.org/10.1002/bse.2427.

Rao, K. K., C. A. Tilt, and L. H. Lester. 2012. “Corporate Governance and Environmental Reporting: An Australian Study.” Corporate Governance: The International Journal of Business in Society 12(2): 143–63. https://doi.org/10.1108/14720701211214052.

Rao, K. K., and C. A. Tilt. 2016. “Board Diversity and CSR Reporting: An Australian Study.” Meditari Accounting Research 24(2): 182–210. https://doi.org/10.1108/MEDAR-08-2015-0052.

Roberts, R. W. 1992. “Determinants of Corporate Social Responsibility Disclosure: An Application of Stakeholder Theory.” Accounting, Organizations and Society 17(6): 595–612. https://doi.org/10.1016/0361-3682(92)90015-K.

Saeed, A., H. Riaz, T. A. Liedong, and T. Rajwani. 2022. "The Impact of TMT Gender Diversity on Corporate Environmental Strategy in Emerging Economies." Journal of Business Research 141: 536–551. https://doi.org/10.1016/j.jbusres.2021.11.057.

Sahu, M., W. M. Alahdal, D. K. Pandey, S. R. Baatwah, and M. S. Bajaher. 2025. “Board Gender Diversity and Firm Performance: Unveiling the ESG Effect.” Sustainable Futures 9: 100493. https://doi.org/10.1016/j.sftr.2025.100493.

Sandretto, D., A. Rizzi, and G. Esposito. 2025. "Gender Diversity Leadership and ESG Performance: The Influence of Women on Boards and in Management." Business Strategy and the Environment 34, no. 4: 5075–5094. https://doi.org/10.1002/bse.4241.

Satterthwaite, F. E. 1946. “An Approximate Distribution of Estimates of Variance Components.” Biometrics Bulletin 2(6): 110–14. https://doi.org/10.2307/3002019.

Schütze, F., and B. Sandbaek. 2025. “Taxonomy Disclosure in the EU – A Useful Framework, despite Current Challenges.” The Economists’ Voice 22(1): 161–72. https://doi.org/10.1515/ev-2025-0016.

Schütze, F., and J. Stede. 2024. “The EU Sustainable Finance Taxonomy and Its Contribution to Climate Neutrality.” Journal of Sustainable Finance & Investment 14(1): 128–60. https://doi.org/10.1080/20430795.2021.2006129.

Simpson, W., D. A. Carter, and F. D’Souza. 2010. “What Do We Know about Women on Boards?” Journal of Applied Finance 20(2): 27–39.

Snider, J., R. P. Hill, and D. Martin. 2003. “Corporate Social Responsibility in the 21st Century: A View from the World’s Most Successful Firms.” Journal of Business Ethics 48(2): 175–87. https://doi.org/10.1023/B:BUSI.0000004606.29523.db.

Tamimi, N., and R. Sebastianelli. 2017. “Transparency among S&P 500 Companies: An Analysis of ESG Disclosure.” Management Decision 55(8): 1660–80. https://doi.org/10.1108/MD-01-2017-0018.

Terry, S. J., T.-M. Whited, and A. A. Zakolyukina. 2023. “Information versus Investment.” The Review of Financial Studies 36(3): 1148–91. https://doi.org/10.1093/rfs/hhac052.

Tremblay, M., Y. Gendron, and B. Malsch. 2016. “Gender on Board: Deconstructing the ‘Legitimate’ Female Director.” Accounting, Auditing & Accountability Journal 29(1): 165–90. https://doi.org/10.1108/AAAJ-05-2014-1711.

United Nations. 2015. Transforming Our World: The 2030 Agenda for Sustainable Development. United Nations. https://sdgs.un.org/publications/transforming-our-world-2030-agenda-sustainable-development-17981.

Valls Martínez, M., S. C. Rambaud, and I. M. Parra Oller. 2019. “Gender Policies on Board of Directors and Sustainable Development.” Corporate Social Responsibility and Environmental Management 26(6): 1539–53. https://doi.org/10.1002/csr.1825.

Velte, P. 2016. “Women on Management Board and ESG Performance.” Journal of Global Responsibility 7(1): 98–109. https://doi.org/10.1108/JGR-01-2016-0001.

Velte, P. 2022. “Does Sustainable Corporate Governance Have an Impact on Materiality Disclosure Quality in Integrated Reporting? International Evidence.” Sustainable Development 30(6): 1655–70. https://doi.org/10.1002/sd.2333.

Velte, P. 2025. “Board Gender Diversity and Corporate Social Responsibility: A Literature Review on Critical Mass Theory and Research Recommendations.” Sustainable Development 33(S1): 1233–48. https://doi.org/10.1002/sd.70071.

Wang, J., and B. S. Coffey. 1992. “Board Composition and Corporate Philanthropy.” Journal of Business Ethics 11(10): 771–78. https://doi.org/10.1007/BF00872309.

Wasiuzzaman, S., and W. M. Wan Mohammad. 2020. “Board Gender Diversity and Transparency of Environmental, Social and Governance Disclosure: Evidence from Malaysia.” Managerial and Decision Economics 41(1): 145–56. https://doi.org/10.1002/mde.3099.

World Commission on Environment and Development. 1987. Report of the World Commission on Environment and Development: Our Common Future (Brundtland Report). https://www.are.admin.ch/dam/are/en/dokumente/nachhaltige_entwicklung/dokumente/bericht/our_common_futurebrundtlandreport1987.pdf.download.pdf/our_common_futurebrundtlandreport1987.pdf.

World Resources Institute and World Business Council for Sustainable Development. 2011. Corporate Value Chain (Scope 3) Accounting and Reporting Standard. Greenhouse Gas Protocol. https://ghgprotocol.org/corporate-value-chain-scope-3-standard.

Yang, T., A. D. Huang, and L. Shang. 2025. "Diverse Links: Female Board Representation and ESG in Varied Business Landscapes." Journal of Cleaner Production 490: 144720. https://doi.org/10.1016/j.jclepro.2025.144720.

Yuriev, A., and O. Boiral. 2024. "Sustainability in the Boardroom: A Systematic Review." Journal of Cleaner Production 442: 141187. https://doi.org/10.1016/j.jclepro.2024.141187.

Zahid, M., H. U. Rahman, W. Ali, M. Khan, M. Alharthi, M. I. Qureshi, and A. Jan. 2020. "Boardroom Gender Diversity: Implications for Corporate Sustainability Disclosures in Malaysia." Journal of Cleaner Production 244: 118683. https://doi.org/10.1016/j.jclepro.2019.118683.

Zhang, J. Q., H. Zhu, and H.-B. Ding. 2013. "Board Composition and Corporate Social Responsibility: An Empirical Investigation in the Post Sarbanes-Oxley Era." Journal of Business Ethics 114(3): 381–92. https://doi.org/10.1007/s10551-012-1352-0.