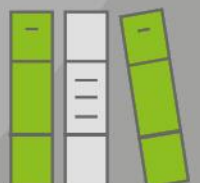


ESG-Ratings: Ansatzpunkte und Aussagegehalt

Dominik Wembacher (B.Sc.)

HFM-Schriftenreihe der
Forschungsstelle Nachhaltigkeit
Ausgabe 1 | 03.05.2024



Impressum

Herausgeber

Forschungsstelle Nachhaltigkeit
Hochschule für Finanzwirtschaft & Management
Simrockstraße 4, 53113 Bonn
www.s-hochschule.de

Verantwortlich

Prof. Dr. Claudia Breuer
Tel.: +49 228 204 9922
E-Mail: claudia.breuer@s-hochschule.de

Redaktion

Matthias Brust (M.Sc.)
Tel.: +49 228 204 9805
E-Mail: matthias.brust@s-hochschule.de

Petra Jonas
Tel.: +49 228 204 9803
E-Mail: petra.jonas@s-hochschule.de

ISSN 2943-5498

Bonn, 03.05.2024

Abstract

Die Nachhaltigkeit von Unternehmen und Investitionsprojekten wird häufig anhand von ESG-Ratings bewertet. Ziel dieses Beitrags ist es, Anbieter und Methoden der ESG-Ratings näher zu beleuchten und den Aussagegehalt von Ratingnoten auf Basis verschiedener wissenschaftlicher Studien kritisch zu beurteilen. Es wird nachgewiesen, dass die ESG-Ratings verschiedener Ratingagenturen für ein und dasselbe Unternehmen teilweise deutlich voneinander abweichen. Als Ursachen werden u. a. fehlende Standards sowie ein abweichendes Vorgehen der Ratingagenturen bei der Rating-Erstellung in Form von Unterschieden bei den verwendeten Attributen, differierenden Messmethoden oder unterschiedlichen Gewichtungen der Dimensionen E (Environmental), S (Social) und G (Governance) identifiziert. Die voneinander abweichenden Ratingnoten sind ein Grund für die herrschende Verunsicherung und stellenweise geringe Akzeptanz bei verschiedenen Investorengruppen. Mangelnde Vergleichbarkeit und Intransparenz erschweren die Bewertung der Nachhaltigkeit von Unternehmen und Investitionsprojekten und damit die Finanzierungs- und Anlageentscheidung unter Berücksichtigung von Nachhaltigkeitskriterien.

Keywords: Corporate Social Responsibility, ESG-Ratings, ESG-Scores, Nachhaltige Investments, Nachhaltigkeitsberichterstattung, Nachhaltigkeitsrisiken, Ratingagenturen

Ich bedanke mich bei Herrn Prof. Dr. Bernd Heitzer, Hochschule für Finanzwirtschaft & Management, und der Forschungsstelle Nachhaltigkeit für wertvolle Kommentare und hilfreiche Anregungen bei der Erstellung dieses Beitrags.

Inhaltsverzeichnis

1	Einleitung	5
2	Definitorische Grundlagen rund um ESG-Ratings	5
3	Vergleich verschiedener Ratingagenturen.....	7
3.1	Marktüberblick.....	7
3.2	Überblick bekannter ESG-Ratingagenturen.....	7
4	Analyse des Aussagegehalts der ESG-Ratingergebnisse.....	9
4.1	Belegung von Divergenzen zwischen Ratingergebnissen.....	9
4.2	Gründe für mangelnde Vergleichbarkeit	11
4.3	Muster bei ESG-Ratings.....	16
5	Zusammenfassung und Ausblick	17
	Literaturverzeichnis	19

1 Einleitung

Vor dem Hintergrund der zunehmenden Bedeutung des Themas Nachhaltigkeit und einem exponentiellen Wachstum nachhaltiger Geldanlagen stellt sich zwangsläufig die Frage, wer Nachhaltigkeit definiert und entscheidet, welche Unternehmen oder Projekte nachhaltig sind und welche nicht. Unternehmen, die als „nachhaltig“ klassifiziert werden, können Finanzierungsvorteile erlangen, z. B. durch Aufnahme in Publikumsfonds mit Nachhaltigkeitsschwerpunkt. Häufig werden diese Entscheidungen auf der Grundlage von ESG-Ratings getroffen. Ein gutes ESG-Rating fungiert dabei als eine Art Gütesiegel, welches den Anlegerinnen und Anlegern signalisieren soll, dass mit ihrem Investment sozial und ökologisch verantwortungsvoll handelnde Unternehmen gefördert werden und sie so einen Beitrag zu einer besseren, nachhaltigeren Zukunft leisten. Ziel dieses Beitrags ist daher eine kritische Beurteilung des Aussagegehaltes und der Validität von ESG-Ratings. Nach einer kurzen Einführung zu den Grundlagen von ESG-Ratings wird ein Überblick über die wichtigsten Anbieter gegeben. Anschließend erfolgt auf Basis verschiedener Untersuchungen der Nachweis, dass die ESG-Ratingnoten teils erheblich voneinander abweichen. Zentraler Gegenstand dieses Beitrags ist eine Analyse der Ursachen der Abweichungen und eine Bewertung der aktuellen Validität von ESG-Ratingnoten zur Verwendung als Kriterium für nachhaltige Finanzierungsentscheidungen.

2 Definitorische Grundlagen rund um ESG-Ratings

Die *UN Global Compact Initiative* hat mit ihrem Bericht „Who Cares Wins“ im Jahre 2004 erstmals den Begriff ESG im Finanzmarktumfeld geprägt.¹ Über die Jahre bezogen sich diverse Veröffentlichungen immer wieder auf diesen Bericht und verfeinerten die Definitionen zu ESG und ESG-Risiken, sodass es heute einen breiten Konsens bezüglich der Bedeutung und Auslegung gibt.

Die *Bundesanstalt für Finanzdienstleistungsaufsicht (BaFin)* hat im Dezember 2019 ein „Merkblatt zum Umgang mit Nachhaltigkeitsrisiken“ veröffentlicht, in dem sie die ESG-Risiken noch einmal genauer definiert. Dabei steht das E in ESG für Environmental und beschreibt Faktoren wie den Klimawandel oder den Umgang mit natürlichen Ressourcen. Das S steht für Social und beinhaltet etwa die Einhaltung anerkannter arbeitsrechtlicher Standards oder eine angemessene Entlohnung. Die dritte Dimension wird durch G für Governance abgedeckt und umfasst Faktoren wie etwa Steuerehrlichkeit oder die Verhinderung von Korruption. Darüber hinaus unterteilt das Merkblatt Risiken für Unternehmen in physische und transitorische Risiken: Physische Risiken sind zum Beispiel Extremwetterereignisse, Transitionsrisiken entstehen dagegen aus der Veränderung der Wirtschaft, etwa mit dem Verbot von Verbrennungsmotoren.²

Für ESG-Ratings gibt es dagegen aktuell noch keine derart vereinheitlichte Definition. Die *Europäische Wertpapier- und Marktaufsichtsbehörde (ESMA)* hat in einem Schreiben an die *Europäische Kommission* eine noch recht weit gefasste Definition vorgeschlagen, die in ESG-Ratings eine Meinung zu den Auswirkungen eines Unternehmens, Emittenten oder Schuldtitels auf ESG-Faktoren, die Anpassung an internationale Klimavereinbarungen oder Nachhaltigkeitsmerkmalen sieht, die anhand eines definierten Rankingsystems abgegeben werden.³ Zudem muss der Verwender oder die Verwenderin vor jeder Betrachtung eines ESG-Ratings prüfen, über welche Perspektive der Risikobetrachtung das Rating genau Auskunft erteilt. So gibt es ESG-Ratingagenturen, die eine Inside-Out-Betrachtung vornehmen, also untersuchen, welchen Einfluss das unternehmerische Han-

¹ Vgl. Vereinte Nationen (Hrsg.) (2004), S. 2.

² Vgl. Bundesanstalt für Finanzdienstleistungsaufsicht (Hrsg.) (2020), S. 13.

³ Vgl. ESMA (Hrsg.) (2021), S. 106.

deln auf die Umwelt hat, während andere prüfen, wie sich Veränderungen in der Umwelt und Gesellschaft auf das Unternehmen auswirken (Outside-In-Betrachtung), sowie Mischformen beider Ansätze.⁴

Im Kreditgeschäft stellen Ratings bereits ein etabliertes Instrument dar, um vor Geschäftsabschluss das Investment schnell bewerten und einsortieren zu können. Es liegt daher nahe, dieses System auch auf Investmententscheidungen unter dem Gesichtspunkt der Nachhaltigkeit zu übertragen. Dabei unterscheidet sich der Aussagegehalt der zwei Ratingarten erheblich und das Ergebnis ist diesbezüglich unterschiedlich einzuordnen. Grundsätzlich gilt, dass Kreditratingagenturen Aussagen zur zukünftigen Zahlungsfähigkeit, abgeleitet aus den Finanzkennzahlen und -daten wie der Finanzierungsstruktur oder Ertragszahlen, treffen. Dagegen beziffern ESG-Ratingagenturen Chancen, aber vor allem auch Risiken im Zusammenhang mit nachhaltigen Entwicklungen und beleuchten im Speziellen die Umweltrisiken sowie soziale und langfristige ökonomische Risiken für ein Unternehmen.⁵

Neben der Zielsetzung gibt es weitere wichtige Unterschiede zwischen ESG- und Kreditratings. Zum einen ist die Kreditwürdigkeit klar als Ausfallwahrscheinlichkeit definiert, während ESG-Leistungen weniger einheitlich definiert sind. ESG-Ratings sind ein Konzept, dessen Grundwerte vielfältig sind und sich noch stetig weiterentwickeln. Des Weiteren sind die Standards bei Kreditwürdigkeitsprüfungen aus der Finanzberichterstattung der Unternehmen über Jahrzehnte gereift und haben sich über die Zeit angeglichen. Die Nachhaltigkeitsberichterstattung ist dagegen noch jung und oftmals freiwillig, einheitliche Standards und Offenlegungsrichtlinien fehlen oder konkurrieren zueinander.⁶

Während für die Refinanzierung Kreditratings bereits obligatorisch sind, kann sich für Unternehmen der Erhalt eines ESG-Ratings mehrfach lohnen. So erhalten Unternehmen, die in ESG-Indizes aufgenommen werden, mehr Aufmerksamkeit und werden von Anlegern Anlegerinnen bereits durch die Aufnahme in einen Index nach ESG-Kriterien als nachhaltig empfunden.⁷ Ein nicht vorhandenes ESG-Rating kann sich dagegen negativ auf den Unternehmenswert auswirken, da das Fehlen eines Ratings möglicherweise als mangelnde Transparenz interpretiert wird.⁸ Des Weiteren fanden Untersuchungen heraus, dass eine Erhöhung des durchschnittlichen ESG-Ratings bei Unternehmen zu sinkenden Kapitalkosten führt. Dabei ist der Effekt auf die Eigenkapitalkosten höher als auf die Fremdkapitalkosten.⁹

Für Investorinnen und Investoren geht es bei ESG-Ratings vorrangig darum, das derzeit noch nicht gemanagte ESG-Risiko eines Unternehmens einordnen zu können und die Informationsasymmetrie zwischen ihnen und dem Unternehmen auf diese Weise zu verringern.¹⁰ ESG-Ratings sollen aber nicht nur Risiken aufzeigen, sondern viele Anlegerinnen und Anleger sind der Meinung, dass ESG-Informationen für die finanzielle Anlageperformance relevant sind und Renditechancen erkennen lassen.

⁴ Vgl. Da Graca, A. (2022), S. 162.

⁵ Vgl. Ossenbrink, N. (2019), S. 131-132.

⁶ Vgl. Berg, F./ Kölbel, J./ Rigobon R. (2022), S. 1319-1320.

⁷ Vgl. Horn, M. (2023), S. 4.

⁸ Vgl. Horn, M. (2023), S. 1.

⁹ Vgl. Bams, D./ van der Kroft, B. (2022), S. 15.

¹⁰ Vgl. Horn, M. (2023), S. 1; Vgl. Amel-Zadeh, A./ Serafeim, G. (2018), S. 89.

3 Vergleich verschiedener Ratingagenturen

3.1 Marktüberblick

Im heutigen Marktumfeld achten bereits viele Kapitalgebende früh auf ESG-Ratings und entscheiden auf deren Grundlage, ob das Unternehmen überhaupt weiter in Betracht gezogen wird. Ein annehmbares ESG-Rating ist somit zu einer Art Mindestanforderung geworden.¹¹ Diese stetig wachsende Bedeutung hat den Markt in den letzten Jahren deutlich verändert. Während längere Zeit ESG-Ratingagenturen ein Nischendasein innehatten, entwickelte sich das Geschäftsfeld der Nachhaltigkeitsratings mit zunehmender Relevanz immer weiter. Neue Akteure kamen auf den Markt, übernahmen bestehende Agenturen oder nutzten ihre vorhandenen Geschäftsstrukturen, um auch in diesem Geschäftsfeld Fuß zu fassen.¹²

Die verschiedenen Ratingagenturen lassen sich dabei in unterschiedliche Gruppen einteilen: Zunächst sind ESG-Ratingagenturen zu nennen, die als Ableger bereits bestehender Kreditratingagenturen zusätzlich ein Nachhaltigkeitsrating anbieten (z. B. *Moody's*, *Fitch* und *S&P*). Daneben agieren aber auch Indexanbieter wie *MSCI* und *FTSE Russell*, die die von ihnen vergebenen ESG-Ratings auch zur Erstellung von ESG-Indizes verwenden. Des Weiteren gibt es Datenplattformen wie *Bloomberg* oder *Refinitiv*, die ihren Kunden als zusätzlichen Service ESG-Ratings anbieten. Neben den bereits genannten existieren jedoch noch eine Vielzahl anderer spezialisierter Firmen und Beratungsunternehmen, die ESG-Ratings erstellen und anbieten.¹³

Generell zeigt sich, dass die Branche stark im Wandel ist und es immer wieder zu Fusionen und Übernahmen kommt, weshalb sich das Umfeld der Ratingagenturen, sowie deren einzelne Ratingprozesse und Ratingbesonderheiten in Zukunft sicher weiterentwickeln und verändern werden.¹⁴ Solange diese Konsolidierung nicht abgeschlossen ist, gibt es aber nach wie vor eine Vielzahl an ESG-Ratingagenturen, die sich um Relevanz und Einfluss bemühen. Allein in der EU bieten derzeit zwischen 30 und 40 ESG-Ratingagenturen ihre Services an.¹⁵ Dazu kommen noch viele andere große Ratinganbieter aus den übrigen Teilen der Welt.

Die einmal vergebenen ESG-Ratingnoten werden von den Agenturen laufend aktualisiert, etwa durch die Aufnahme neuer Informationen, sowie stetiges Screening und Bewertung von neuen Erkenntnissen.¹⁶ ESG-Ratings werden in der Regel jährlich überprüft und überarbeitet, in Ausnahmefällen und nach Auftreten entscheidender neuer Erkenntnisse auch ad hoc.¹⁷ Je nach Ratingzeitpunkt kann hier jedoch bereits der erste Grund für mögliche Ratingdifferenzen festgemacht werden, denn je nach Zeitpunkt des Abschlusses der Datensammlung, können die Ratings auch für das selbe Jahr auf unterschiedliche Datengrundlagen zurückgreifen und so eine abweichende Interpretation der ESG-Leistung erfolgen.

3.2 Überblick bekannter ESG-Ratingagenturen

Der Markt für ESG-Ratings ist sehr heterogen und die Anzahl der Ratingagenturen unübersichtlich. Dennoch gibt es eine Reihe von Ratinganbietern, die sich bereits etablieren konnten und deren Namen vorrangig genannt werden, wenn es um Ratings mit Nachhaltigkeitsausrichtung geht.

¹¹ Vgl. Leinweber, L./ Kögler, A. (2023), S. 65.

¹² Vgl. Escrig-Olmedo, E. et al. (2019), S. 6.

¹³ Vgl. Mazzacurati, J. (2021), S. 108.

¹⁴ Vgl. Escrig-Olmedo, E. et al. (2019), S. 14; Vgl. Mazzacurati, J. (2021), S. 107.

¹⁵ Vgl. Wohnhaas, B. (2023), S. 26.

¹⁶ Vgl. Da Graca, A. (2022), S. 161.

¹⁷ Vgl. Christensen, D./ Serafeim, G./ Sikochi, A. (2022), S. 154; Vgl. Da Graca, A. (2022), S. 161.

Unter diesen sind aktuell *MSCI* und *Sustainalytics*, die derzeit mit zu den einflussreichsten ESG-Ratingagenturen zählen.¹⁸ Ausgewählte Ratingagenturen, die auch in einem Großteil der später erwähnten Studien und Untersuchungen genannt werden, sowie deren Arbeitsweisen werden daher kurz beleuchtet. Die Ausführungen sollen ein Schlaglicht auf die führenden ESG-Ratingagenturen werfen und die Heterogenität unter den Anbietern verdeutlichen.

MSCI veröffentlicht ESG-Bewertungen für 8.500 Unternehmen sowie 14.000 Emittenten weltweit und beschäftigt mehr als 200 Analysten. Die Daten der MSCI-ESG-Research-Analysten werden unter anderem für Portfolio-Analysen und zur Erstellung von ESG-Indizes verwendet.¹⁹ *MSCI* nutzt bei seinen Datenerhebungen nur öffentlich zugängliche Berichte, etwa von Behörden oder den Unternehmen selbst. Die Bewertung erfolgt dabei branchenspezifisch anhand von 37 definierten Schlüsselthemen und die Ergebnisse werden in einem ESG-Gesamtrating dargestellt, dessen Skala von CCC bis AAA reicht. Dabei werden Unternehmen mit AAA und AA als ESG-Führer eingestuft, mit A, BBB und BB als Durchschnitt und B sowie CCC als Nachzügler.²⁰ *MSCI* bewertet dabei mit seinem Rating, wie sehr Unternehmen von finanziell relevanten Nachhaltigkeitsrisiken entlang den drei Dimensionen Umwelt (E), Soziales (S) und Unternehmensführung (G) betroffen sind und wie deren Fähigkeit eingeschätzt wird, diese zu managen.²¹ *MSCI*, das vor allem für die Erstellung von Indizes bekannt ist, erweiterte das Portfolio an Leistungen um ESG-Ratings und wuchs schnell durch eine Reihe von Übernahmen anderer Ratinganbieter, unter anderem *RiskMetrics*, die zuvor selbst zwei namhafte ESG-Ratingagenturen, *KLD* und *Innovest*, übernommen hatte.²²

Eine zweite namhafte ESG-Ratingagentur ist *Sustainalytics*, das ESG-Bewertungen für weltweit mehr als 16.000 Unternehmen bereitstellt und damit nach eigenen Angaben die meisten Ratings im Portfolio hat. *Sustainalytics* verwendet eine Perzentil-Bewertung von 1 bis 100 basierend auf ihrem ESG-Gesamtwert, der Unternehmen mit deren Branche vergleicht. Dabei bescheinigt ein Ergebnis von 1 bis 10 dem Unternehmen ein unwesentliches oder vernachlässigbares ESG-Risiko, ein Ergebnis zwischen 10 und 20 ein geringes, zwischen 20 und 30 ein mittleres, zwischen 30 und 40 ein hohes und über 40 ein ernstes ESG-Risiko.²³ *Sustainalytics* gehört zu *Morningstar*, das *Sustainalytics* 2020 übernommen hat und dessen Hauptgeschäft die Bewertung von Investmentfonds und börsengehandelten Fonds für Privatanleger und -anlegerinnen ist.²⁴ Auf der Basis von *Sustainalytics*-Ratings werden daher auch ESG-Ratings auf Fondsebene gebildet.²⁵

Etwas kleiner ist die Ratingagentur *S&P Global*. Sie stellt aktuell für etwa 8.000 Unternehmen weltweit ESG-Bewertungen zur Verfügung, die zum einen auf Ebene einer Gesamtbewertung sowie in Einzelscores für die Dimensionen E, S und G veröffentlicht werden.²⁶ *S&P Global* bewertet dabei jede Branche mit individuellen Schlüsselmerkmalen und gewichtet zudem die drei Einzeldimensionen nicht gleichmäßig, sondern legt den größten Fokus auf die Governance-Faktoren.²⁷ Auch *S&P Global* verwendet eine Perzentil-Bewertung von 1 bis 100 und ordnet die ESG-Gesamtwerte der Unternehmen im Vergleich zu der jeweiligen Branche ein. Im Gegensatz zu *Sustainalytics* entspricht

¹⁸ Vgl. Lykkesfeldt, P./ Kjaergaard, L. L. (2022), S. 307.

¹⁹ Vgl. MSCI (Hrsg.) (2020), S. 4-5.

²⁰ Vgl. MSCI (Hrsg.) (2020), S. 4-6.

²¹ Vgl. Berg, F./ Heeb, F./ Kölbel, J. (2022), S. 8.

²² Vgl. Avetisyan, E./ Hockerts, K. (2017), S. 323.

²³ Vgl. Sustainalytics (Hrsg.) (2023).

²⁴ Vgl. Larcker, D. F. et al. (2022), S. 3.

²⁵ Vgl. Christensen, D./ Serafeim, G./ Sikochi, A. (2022), S. 154.

²⁶ Vgl. S&P Global (Hrsg.) (2023a).

²⁷ Vgl. S&P Global (Hrsg.) (2023b), S. 5-6.

jedoch ein Wert von 100 einer sehr guten ESG-Bewertung, während 1 der niedrigste zu erreichende Wert ist.²⁸

Eine vierte bedeutende Agentur ist *Refinitiv*, die ESG-Bewertungen für etwa 12.500 Emittenten weltweit veröffentlicht. Dabei erfasst *Refinitiv* mehr als 630 ESG-Kennzahlen je Unternehmen und bezieht die wichtigsten und vergleichbarsten 183 je Branche in die Gesamtbewertung des Scores mit ein. Die ESG-Gesamtbewertung sowie die Teilwertungen der Einzeldimensionen erfolgt über eine Perzentil-Bewertung und liegt zwischen 0 (Minimalwert) und 100 (Maximalwert). 0 bis 25 Punkte entsprechen dabei dem ersten Quartil, das Unternehmen eine schwache ESG-Leistung attestiert, 25 bis 50 Punkte eine zufriedenstellende, 50 bis 75 eine gute und 75 bis 100 Punkte eine im Branchenvergleich sehr gute ESG-Leistung. Die Beurteilung der relativen ESG-Leistung des Unternehmens für E- und S-Faktoren erfolgt im Vergleich mit der Branche, für G-Faktoren gegenüber Unternehmen mit demselben Gründungsland.²⁹

Nachfolgend zeigt eine eigens angefertigte Gegenüberstellung die teilweise signifikanten Bewertungs- und Darstellungsunterschiede für ausgewählte Unternehmen.

Rating-agentur Bewertetes Unternehmen	MSCI	Sustainalytics	S&P Global	Refinitiv
Microsoft Corp.	AAA	15,2	55	88
SAP SE	AAA	12,6	74	89
Tesla, Inc.	A	25,3	40	70
Volkswagen AG	B	26,4	55	81
Barclays PLC	AA	23,8	59	81
Shell PLC	AA	33,7	46	93

Die Ratingnoten stammen von den Internetseiten der jeweiligen Ratingagenturen selbst (Stand: 11.02.2024).³⁰

4 Analyse des Aussagegehalts der ESG-Ratingergebnisse

4.1 Belegung von Divergenzen zwischen Ratingergebnissen

Als erste Kennzahl wird der Korrelationskoeffizient zwischen den Ratingergebnissen verschiedener ESG-Ratingagenturen für dieselben Unternehmen herangezogen. Eine Studie von *Berg, Kölbel* und *Rigobon* (2022) untersuchte die Ratingergebnisse zwischen den Ratingagenturen *Sustainalytics*, *Moody's ESG*, *S&P Global*, *KLD*, *Refinitiv* und *MSCI*. Dabei kamen sie zu dem Ergebnis, dass obwohl es sich bei allen um große, globale ESG-Ratingagenturen handelt, die die Unternehmen umfassend beurteilen, die Korrelationswerte lediglich zwischen 0,38 und 0,71 liegen.³¹

²⁸ Vgl. S&P Global (Hrsg.) (2023b), S. 10.

²⁹ Vgl. Refinitiv (Hrsg.) (2022), S. 3-6.

³⁰ Vgl. MSCI (Hrsg.) (2024); Vgl. Sustainalytics (Hrsg.) (2024); Vgl. S&P Global (Hrsg.) (2024); Vgl. Refinitiv (Hrsg.) (2024).

³¹ Vgl. Berg, F./ Kölbel, J./ Rigobon R. (2022), S. 1316.

Die höchste Übereinstimmung zueinander haben in der Untersuchung dabei die Ratings von *Sustainalytics* und *Moody's ESG* mit einem Korrelationswert auf ESG-Gesamtebene von 0,71. Die Ratings von *KLD* (heute *MSCI*) weisen dagegen die geringsten Korrelationen zu den anderen Ratingagenturen auf, sowohl auf ESG-Gesamtratingebene als auch auf der Ebene der Einzeldimensionen Environment, Social und Governance. Bei allen betrachteten ESG-Ratingagenturen hat die Umwelt-dimension im Unternehmensvergleich in der Studie von *Berg, Kölbel und Rigobon (2022)* die höchste Korrelation, mit einem Durchschnittswert von 0,53, gefolgt von der Social-Dimension mit einem Wert von 0,42. Am geringsten korrelieren die Ergebnisse der Agenturen bei der Einzelbetrachtung auf Governance-Ebene. Hier ergibt sich bei der Betrachtung im Durchschnitt lediglich ein Korrelationswert von 0,30.³² Einzelne Paare von ESG-Ratingagenturen haben bei Betrachtung der Korrelation der Ergebnisse für die Governance-Ebene sogar ein negatives Ergebnis von -0,05.³³

Auch *Da Graca (2022)* untersuchte für seine Studie die Korrelation der Ratingergebnisse von vier Ratingagenturen (*MSCI*, *Sustainalytics*, *Refinitiv* und *ISS*) anhand von 2.758 betrachteten Einzelunternehmen. Dabei ergaben sich Korrelationswerte zwischen 0,24 und 0,60.³⁴

Chatterji, Durand, Levine und Touboul (2014) kamen bei ihrer Untersuchung zu einem ähnlichen Ergebnis. Sie fanden jedoch zusätzlich heraus, dass die durchschnittlichen Korrelationen zwischen Ratingagenturen getrennt nach dem Sitz in den USA und dem Sitz in Europa untereinander jeweils höher sind als im Vergleich aller Ratingagenturen herkunftsübergreifend.³⁵ Dies lässt den Schluss zu, dass es je nach geografischer Herkunft verschiedene theoretische Standards gibt, was unter Nachhaltigkeit zu verstehen ist. So lagen die Korrelationen der Ratingergebnisse für Unternehmen in den verschiedenen untersuchten ESG-Indizes insgesamt bei einer durchschnittlichen Gesamtkorrelation von 0,30, die Korrelation für in Europa ansässigen Ratingagenturen lag dagegen höher bei 0,54 und für in den USA beheimateten Agenturen bei 0,47.³⁶

Zur besseren Einordnung der Ergebnisse aus der Betrachtung der Korrelationskoeffizienten kann die Korrelation zwischen Kreditratings von zwei führenden Kreditratingagenturen *Moody's* und *Standard&Poor's* herangezogen werden. Die untersuchten Unternehmens-Kreditratings hatten mit einer Korrelation von durchschnittlich 0,99 eine sehr viel höhere Übereinstimmung.³⁷ Beide Ratingagenturen sind jedoch in den Vereinigten Staaten von Amerika beheimatet, auf eine Untersuchung von unterschiedlichen Korrelationen der Kreditratings von US-amerikanischen zu europäischen Kredit-Ratingagenturen wird hier jedoch verzichtet. Diese deutlich geringeren Korrelationen bei den ESG-Ratings zeigen, wie grundverschieden die ESG-Ratingnoten zueinander sind und, dass Zweifel an deren Aussagekraft in Teilen durchaus berechtigt erscheinen.

Berg, Kölbel und Rigobon (2022) untersuchten die Ratingergebnisse darüber hinaus auch anhand der Berechnung des *Krippendorff's Alpha*³⁸, um die Reliabilität der Ratingnoten der verschiedenen Agenturen zu messen. Während *Krippendorff* einen Wert von größer 0,80 für eine aussagekräftige Verlässlichkeit verlangt und einen Wert von 0,667 als untere Grenze ansieht, kommt die Studie

³² Vgl. Berg, F./ Kölbel, J./ Rigobon R. (2022), S. 1321.

³³ Vgl. Berg, F./ Kölbel, J./ Rigobon R. (2022), S. 1322.

³⁴ Vgl. Da Graca, A. (2022), S. 164.

³⁵ Vgl. Chatterji, A. et al. (2014), S. 13.

³⁶ Vgl. Chatterji, A. et al. (2014), S. 14.

³⁷ Vgl. Da Graca, A. (2022), S. 164.

³⁸ Wenn Beobachter in ihren Ergebnissen vollkommen übereinstimmen, ergibt sich ein Krippendorff-Alpha von 1,0, was auf perfekte Zuverlässigkeit und Übereinstimmung hinweist. Ein Krippendorff-Alpha von 0 bedeutet, dass keine Zuverlässigkeit der Übereinstimmungen angenommen werden kann; Vgl. Tschirschwitz, D. (2022), S. 357-358.

beim Vergleich der sechs Ratingagenturen und den untersuchten 924 Unternehmen auf ein *Krippendorff* Alpha von 0,55. Dies unterstreicht die bereits festgestellte erhebliche Diskrepanz zwischen den verschiedenen ermittelten ESG-Ratings für dieselben Unternehmen und stellt deren Aussagegehalt weiter in Frage.³⁹

Da Graca (2022) betrachtete neben der Korrelation noch die Verteilungskurven der ESG-Ratingnoten der vier untersuchten ESG-Ratingagenturen für die 2.758 untersuchten Unternehmen und stellte fest, dass es keine Normalverteilung gibt. Die visualisierten Ratingergebnisse weichen sogar deutlich voneinander ab. Die Ratings von *Sustainalytics* und *Refinitiv* weisen eine linksschiefe Verteilung auf, was bedeutet, dass deutlich mehr Unternehmen ein gutes bis sehr gutes Ratingergebnis erhalten, als die Ratingagenturen Unternehmen eine schlechte ESG-Leistung attestieren und niedrige Ratingnoten vergeben. Die aggregierten Ratingergebnisse von *ISS ESG* weisen dagegen eine rechtsschiefe Verteilung auf, was aussagt, dass überdurchschnittlich niedrige ESG-Bewertungen, beziehungsweise schlechtere Ratingnoten vergeben wurden. Das Histogramm der Ergebnisse von *MSCI* entspricht dabei noch am ehesten einer Normalverteilung, wobei auch hier eine links-schiefe Tendenz zu erkennen ist.⁴⁰

Christensen, Serafeim und Sikochi (2022) prüften zusätzlich, ob sich die ESG-Ratingergebnisse im Zeitverlauf einander angenähert haben, was aufgrund der fortschreitenden Regulatorik und versuchten Standardisierung durch Behörden erwartbar gewesen wäre und in anderen Geschäftsfeldern auch so eingetreten ist. Das Ergebnis ist jedoch das Gegenteil. Die Studie fand heraus, dass die ESG-Ratingabweichungen im Laufe der Zeit zugenommen haben, obwohl sich infolge der gestiegenen Offenlegungen und verbesserten Nachhaltigkeitsberichterstattung der Unternehmen die durchschnittlichen ESG-Ratings verbessert haben.⁴¹

4.2 Gründe für mangelnde Vergleichbarkeit

Gemäß *Amel-Zadeh und Serafeim* (2018) bemängeln viele Anleger und Anlegerinnen die schwierige Vergleichbarkeit sowie das Fehlen von Standards bei der Veröffentlichung von ESG-Informationen und Ratings.⁴² Gründe für Abweichungen liegen dabei meist im Ratingprozess begründet, wie unterschiedliche Studien herausfanden.

Als erster Grund für Differenzen bei den Ratingergebnissen werden oftmals die fehlenden Standards angeführt. Im Gegensatz zu Finanzinformationen wie Verschuldungsgrad oder Rentabilitäten besteht bei ESG-Informationen noch keine übergreifende Einigkeit darüber, welche genauen Messgrößen herangezogen werden müssen und wie die Bedeutung der Informationen für die ESG-Leistungen ermittelt werden soll oder wie bestimmte Kennzahlen und Informationen interpretiert und beurteilt werden müssen.⁴³

Vor allem auf Gesamt-ESG-Basis gibt es bisher nur wenig allgemeingültige Standards. Der Umweltbereich ist gegenwärtig am stärksten im Fokus der Öffentlichkeit, der sozialen Dimension wird dagegen oftmals weniger Beachtung geschenkt oder wird ganz außen vorgelassen. Generell fehlen klare Vorgaben und Standards für Kennzahlen.⁴⁴

³⁹ Vgl. Berg, F./ Kölbel, J./ Rigobon R. (2022), S. 1322.

⁴⁰ Vgl. Da Graca, A. (2022), S. 163.

⁴¹ Vgl. Christensen, D./ Serafeim, G./ Sikochi, A. (2022), S. 165.

⁴² Vgl. Amel-Zadeh, A./ Serafeim, G. (2018), S. 97.

⁴³ Vgl. Christensen, D./ Serafeim, G./ Sikochi, A. (2022), S. 158.

⁴⁴ Vgl. Leker, J./ Frieden, F. (2022), M1.

Neben den differierenden Auffassungen, wie nachhaltiges Handeln zu interpretieren ist, liegt eine weitere Schwierigkeit in der Verwendung von öffentlichen Informationen für die Erstellung eines Ratings. Denn die sehr unterschiedliche Art und Weise, wie Unternehmen Informationen und Daten offenlegen und kommunizieren sowie die abweichende Verfügbarkeit und Aufbereitung dieser Informationen erschwert die Vergleichbarkeit, was sich am Ende auch in einem abweichendem Rating-Ergebnis widerspiegeln kann.⁴⁵ Zu den fehlenden einheitlichen, internationalen Standards für Nachhaltigkeitsberichterstattung kommen darüber hinaus zusätzlich oft noch weitere nationale und branchenspezifische Vorgaben, die die Unterschiede bei den Datengrundlagen der Ratingagenturen über das zu bewertende Unternehmen noch verstärken.⁴⁶

Ferner fehlt es an Transparenz, was den Ratingprozess angeht und einen Ergebnisvergleich weiter erschwert. Aufgrund dieses Black-Box Problems muss sich die Branche der ESG-Ratingagenturen den Vorwurf der Intransparenz und undurchsichtigen Komplexität ihrer Arbeit aktuell noch gefallen lassen.⁴⁷

Dennoch ist anzumerken, dass auch wenn aktuell noch keine einheitlichen Vorgehensweisen und bindenden Vorgaben zur Erstellung eines Ratings existieren, es doch auf Unternehmensebene bereits Standards gibt, an denen sich die Agenturen bei der Bewertung der Kriterien für die einzelnen Unternehmen orientieren können. Hier gibt es neben den Leitsätzen für multinationale Unternehmen der Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD), den UN Global Compact und den Empfehlungen der European Foundation for Quality Management für die S- und G-Säule auch weitere Empfehlungen und Richtlinien für die E-Säule, wie die des EMAS (Eco Management and Audit Scheme), die Global Reporting Initiative oder die Umweltnormreihe ISO14000ff für Fragen zum Umweltmanagement und der damit verbundenen Berichterstattung.⁴⁸ Dies zeigt, dass die Branche stetig im Wandel ist und sich in Zukunft weitere Standards und einheitliche Grundsätze entwickeln werden. Wie sich diese auf die bestehenden Kritikpunkte auswirken, bleibt abzuwarten.

Eng mit den fehlenden Standards ist auch die Kritik an den verwendeten Attributen und Daten verbunden. Denn neben der Ratingkonzeptionen bleibt der Öffentlichkeit meist auch in großen Teilen verborgen, welche Daten tatsächlich verwendet werden.⁴⁹ Berg, Kölbel und Rigobon (2022) untersuchten die Abweichungen von ESG-Ratings verschiedener Anbieter und konzentrierten sich dabei auf verschiedene Abweichungsgründe. Dabei kamen sie zu dem Ergebnis, dass 38 % der Abweichungen auf den Umfang der Daten und Attribute zurückzuführen sind.⁵⁰

Die Ratingagenturen greifen für ihren Ratingprozess auf verschiedene verfügbare Daten zurück, wie etwa Unternehmensdaten und Informationen von Behörden oder von den Unternehmen selbst. Darüber hinaus berücksichtigen sie auch die Auswertungen von sozialen Initiativen, die Beziehung zu Investoren und Investorinnen, zur Kundschaft oder zu den Mitarbeitenden sowie kontroverse Geschäftstätigkeiten oder den Umgang mit der Umwelt. Gerade das Umweltmanagement beziehen alle Ratingagenturen mit ein, während Kriterien wie etwa der Schutz der Rechte indigener Völker oder die Lobbyarbeit lediglich von einzelnen Agenturen berücksichtigt werden.⁵¹

⁴⁵ Vgl. Windolph, S. E. (2011), S. 67; Vgl. Leker, J./ Frieden, F. (2022), M1.

⁴⁶ Vgl. Schäfer, H. (2020), S. 379-380.

⁴⁷ Vgl. Schäfer, H. (2020), S. 377-378.

⁴⁸ Vgl. Escrig-Olmedo, E./ Muñoz-Torres, M. J./ Fernández-Izquierdo, M. A./ (2010), S. 452.

⁴⁹ Vgl. Leker, J./ Frieden, F. (2022), M1.

⁵⁰ Vgl. Berg, F./ Kölbel, J./ Rigobon R. (2022), S. 1317.

⁵¹ Vgl. Escrig-Olmedo, E./ Muñoz-Torres, M. J./ Fernández-Izquierdo, M. A./ (2010), S. 447; Vgl. Berg, F./ Kölbel, J./ Rigobon R. (2022), S. 1317.

Auch wenn alle Agenturen dabei das Ziel verfolgen, eine objektive, unabhängige und letztlich auch nachvollziehbare Bewertung abzugeben, so verwendet jede Ratingagentur dafür unterschiedliche Methoden, um ihre Daten zu erhalten. Mögliche Datenquellen sind etwa der Internetauftritt der Unternehmen, die Analyse von Medienberichten der Unternehmen und nicht zuletzt die aktive Kommunikation mit dem zu bewertenden Unternehmen etwa durch Fragebögen und Interviews. Dabei verwenden manche Ratingagenturen alle Datenquellen, während andere nur die öffentlich zugänglichen verwenden. Einzelne ESG-Ratingagenturen stehen dagegen meist über den gesamten Ratingprozess in engem Kontakt mit den Unternehmen. Zudem werden von manchen nur nicht-finanzielle Daten genutzt, um die ESG-Leistung unabhängig von der finanziellen Leistungsfähigkeit des Unternehmens zu messen, während andere nicht-finanzielle und finanzielle Daten kombinieren.⁵²

Neben den verwendeten Datenquellen gibt es bei den verwendeten Attributen noch weitere Diskrepanzen. Viele Indikatoren werden nur von einzelnen Ratingagenturen verwendet, während andere bestimmte Datenpunkte nicht in ihrem Ratingprozess einbeziehen. So versuchten Berg, Kölbel und Rigobon (2022) in ihrer Untersuchung alle verwendeten Datenpunkte einer der drei Kategorien E, S und G zuzuordnen und stellten dabei große Unterschiede fest. Dies wird in der Studie dahingehend ausgelegt, dass es bereits grundsätzlich akzeptierte ESG-Kernelemente gibt, die jede Ratingagentur einbezieht, während andere Kategorien nur von einzelnen Agenturen berücksichtigt werden. Allgemein anerkannte Datenkategorien sind hier z. B. die Biodiversität, Arbeitspraktiken, Vergütungsregelungen oder Lieferketten. Uneinigkeit besteht dagegen bei Fragen zu Steuerehrlichkeit oder der Bewertung von Investitionen der Unternehmen.⁵³

Dabei ist festzuhalten, dass durch verschärfte Offenlegungspflichten die Anzahl der verfügbaren Daten zunimmt. Christensen, Serafeim und Sikochi (2022) fanden dazu heraus, dass eine größere Offenlegung von ESG-Daten, ob freiwillig oder verpflichtend, allerdings zu größeren Abweichungen und Uneinigkeiten bei den ESG-Ratingergebnissen führt.⁵⁴ Daneben hat die stärkere Offenlegung aber auch den Effekt, dass sie bei den Unternehmen zu tendenziell verbesserten Ratings führt. Sie erklären dies damit, dass die erhöhte Offenlegung generell positiv gewertet wird, die Ratingagenturen die zusätzlichen Informationen jedoch sehr individuell interpretieren, was somit Abweichungen verstärkt. Die sich vergrößernden Ratingabweichungen scheinen vor allem bei Unternehmen mit einem schlechteren ESG-Rating verstärkt aufzutreten.⁵⁵

Als Sonderfaktor der verwendeten Attribute kommen fehlende Werte und Zukunftsversprechen eine besondere Bedeutung zu. So stellen Forschende im Zusammenhang mit der Erstellung von ESG-Ratings die Frage, inwieweit geringe Antwortraten auf Umfragen und Fragebögen von einzelnen Agenturen die Verlässlichkeit der Antworten beeinflusst und zu Verzerrungen bei der Bewertung der Unternehmen führt.⁵⁶ Schließlich muss jede Ratingagentur entscheiden, wie sie mit fehlenden Daten umgeht oder wie sie es schafft, aus unterschiedlichen Formen der Nachhaltigkeitsberichterstattung verschiedener Unternehmen standardisierte Variablen für ihr Rating herauszuziehen.⁵⁷ Für das Ratingergebnis ist es daher von entscheidender Bedeutung, wie mit fehlenden

⁵² Vgl. Scalet, S./ Kelly, T. F. (2010), S. 71; Vgl. Da Graca, A. (2022), S. 160 – 161; Vgl. Escrig-Olmedo, E. et al. (2019), S. 3; Vgl. Billio, M. et al. (2021), S. 1429-1431.

⁵³ Vgl. Berg, F./ Kölbel, J./ Rigobon R. (2022), S. 1323-1325.

⁵⁴ Vgl. Christensen, D./ Serafeim, G./ Sikochi, A. (2022), S. 150.

⁵⁵ Vgl. Christensen, D./ Serafeim, G./ Sikochi, A. (2022), S. 162.

⁵⁶ Vgl. Chatterji, A./ Levine, D. (2006), S. 43.

⁵⁷ Vgl. Larcker, D. F. et al. (2022), S. 4.

Daten umgegangen wird. Dabei gibt es etwa die Möglichkeit, fehlende Datenpunkte im Ratingprozess wegzulassen, was allerdings gerade im Branchenvergleich die Vergleichbarkeit der Ratings von Unternehmen, die Werte melden, zu welchen die es nicht tun, negativ beeinflusst.⁵⁸

Ratingagenturen versuchen dieses Problem etwa durch das Treffen von Annahmen zu lösen. Eine der Annahmen kann sein, dass bei fehlenden Informationen oder ausbleibender Offenlegung für ein Attribut, welches in der Branche aber als wichtig angesehen wird, dies in der Folge mit dem schlechtesten Wert belegt wird, um dem Unternehmen eine schlechte ESG-Leistung zu attestieren.⁵⁹ Die danach handelnden Agenturen argumentieren, dass mit fehlenden Informationen negative Daten vor der Öffentlichkeit verheimlicht werden sollen und bestrafen sie daher mit einer sehr vorsichtigen Annahme.⁶⁰ Haben dagegen die meisten Unternehmen einer Branche keine Informationen zu einem Attribut offengelegt, kann die Annahme getroffen werden, dass dieses Kriterium für die Branche von untergeordneter Bedeutung ist und der fehlende Wert deshalb mit dem Branchendurchschnitt ergänzt werden kann.⁶¹

Ähnlich differiert der Umgang mit Plänen und Zielsetzungen für die Zukunft. Während manche Ratingagenturen diese Ausführungen weitestgehend außen vorlassen, zeigt es sich, dass oftmals bereits der Plan zu mehr und konkreteren nachhaltigeren Leistungen das Rating positiv beeinflussen kann, obwohl es teilweise Jahre dauert, bis die Maßnahmen tatsächlich umgesetzt und realisiert werden.⁶² Problematisch erscheint hier besonders, dass die Nachhaltigkeitsberichte von den Unternehmen selbst verfasst werden und so leicht geschönte Versprechen auf eine Verbesserung der zukünftigen Nachhaltigkeitsleistung beinhalten können.

Bei ihrer weiteren Untersuchung der ESG-Ratings konnten Berg, Kölbel und Rigbon (2022) in ihrer Studie zudem 56 % der Abweichungen auf Unterschiede in der Messung zurückführen.⁶³ Mehr als die Hälfte der Ratingdivergenzen kommen also daher, dass ESG-Ratingagenturen bei Ihren Untersuchungen unterschiedliche Messmethoden anwenden, etwa indem sie aus dem Pool an Daten unterschiedliche Kriterien als relevant heranziehen. Bei den genannten 56 % handelt es sich um Durchschnittswerte, die beim Vergleich verschiedener Ratingagenturen schwanken. So sind Differenzen zwischen den Ratingnoten von Moody's ESG und Refinitiv zu 78 % auf die unterschiedlichen Messmethoden zurückzuführen.⁶⁴ Die Rating-Unterschiede, die auf die Messung zurückzuführen sind, ergeben sich auch daraus, dass sie zwar ähnliche Indikatoren verwenden, manche Agenturen aber z. B. unethisches Verhalten der Umwelt- oder Social-Säule zuordnen, während andere dies auf die Unternehmensführung schieben und daher unter Governance berücksichtigen.⁶⁵ Dies zeigt, wie stark die Methoden der verschiedenen Ratingagenturen voneinander abweichen.

Zudem können Divergenzen bei der Messung auftreten, indem zwar dieselben Attribute in den Ratingprozess einfließen, jedoch mit verschiedenen Indikatoren und Datenpunkten gemessen werden. Etwa wenn das Attribut Arbeitspraktiken der Unternehmen einmal anhand der Mitarbeiterfluktuation und einmal anhand der arbeitsrechtlichen Gerichtsprozesse, in die das Unternehmen verwickelt ist, gemessen wird.⁶⁶ Dazu kommt, dass einige Ratingagenturen neben den drei großen

⁵⁸ Vgl. Larcker, D. F. et al. (2022), S. 5.

⁵⁹ Vgl. Christensen, D./ Serafeim, G./ Sikochi, A. (2022), S. 157.

⁶⁰ Vgl. Da Graca, A. (2022), S. 165.

⁶¹ Vgl. Christensen, D./ Serafeim, G./ Sikochi, A. (2022), S. 157; Vgl. Da Graca, A. (2022), S. 165.

⁶² Vgl. Bams, D./ van der Kroft, B. (2022), S. 11.

⁶³ Vgl. Berg, F./ Kölbel, J./ Rigobon R. (2022), S. 1317.

⁶⁴ Vgl. Berg, F./ Kölbel, J./ Rigobon R. (2022), S. 1335-1336.

⁶⁵ Vgl. Lykkesfeldt, P./ Kjaergaard, L. L. (2022), S. 307.

⁶⁶ Vgl. Berg, F./ Kölbel, J./ Rigobon R. (2022), S. 1317.

Säulen E, S und G, wie es die meisten ESG-Ratingagenturen machen, auch verschiedene weitere Dimensionen bei ihrer Raterstellung berücksichtigen. Manche Ratingagenturen nehmen wirtschaftliche Dimensionen mit ins Rating auf, während andere stärker auf die Produkte der Unternehmen eingehen oder die Säulen Social und Governance zu einer zusammenfassen.⁶⁷ Dies zeigt, dass es auch unter den Ratingagenturen verschiedene Spezialisierungen gibt. Mit einer weiteren Konsolidierung der Branche wird es allerdings auch im Feld der Messmethoden aktuell und zukünftig viel Bewegung geben.

Die verschiedenen Messmethoden führen zu großen Abweichungen unter den Ratingagenturen, bis hin zu negativen Korrelationen bei den Ratingergebnissen unter Betrachtung der Messmethoden. Auf der Ebene der Umweltpolitik der Unternehmen konnten noch hohe Korrelationen von durchschnittlich 0,55 nachgewiesen werden. Für andere Themenbereiche sehen die Ergebnisse jedoch anders aus. So konnten bei Indikatorengruppen wie dem Lobbying zwischen den Agenturen Sustainalytics und Moody's ESG oder dem Umgang mit den Rechten indigener Völker bei Sustainalytics und Refinitiv leicht negative Korrelationen nachgewiesen werden⁶⁸, was bedeutet, dass die Ratingagenturen hier bei ihrer Beurteilung zu Ergebnissen kommen, die unabhängig voneinander, bis hin zu leicht gegenläufig sind.

Als weitere Ursachenkategorie für Abweichungen zwischen ESG-Ratings konnte die Gewichtung, nicht nur im Gesamtrating auf Grund der Einzelsäulen E, S, und G, sondern auch anhand der Gewichtung von Themenschwerpunkten und Attributen innerhalb der Säulen festgemacht werden. Wobei die Gewichtung meist einen recht geringen Beitrag leistet. Berg, Kölbel und Rigbon (2022) konnten bei ihren Untersuchungen 6 % der Divergenzen auf differierende Gewichtungen zurückführen.⁶⁹ Auch hier handelt es sich um Durchschnittswerte, sodass es zwischen den einzelnen Ratingagenturen in der Einzelbetrachtung auch zu größeren Abweichungen kommen kann. Etwa bei den Ratings von Sustainalytics und Refinitiv konnte die Gewichtung für 22 % der Abweichung verantwortlich gemacht werden.⁷⁰

Die Differenzen entstehen, da ein Anbieter alle drei ESG-Säulen gleich gewichtet, ein anderer nimmt eine unausgeglichene Gewichtung vor, die in seinen Augen eine angemessenere Berücksichtigung der Einflussfaktoren auf die Unternehmen einer bestimmten Branche widerspiegelt.⁷¹ Des Weiteren haben Untersuchungen herausgefunden, dass sich die Kriterien innerhalb der untersuchten Dimensionen und Säulen verändern. Während 2008 in der Säule Environment noch die Analyse Kriterien Umweltpolitik, Emissionen und Klimawandel als Ganzes von großer Bedeutung im Bewertungsprozess waren, rückten 2018 neben der Umweltpolitik die Wassernutzung und das Wassermanagement sowie der Schutz der Biodiversität in den Fokus der Beurteilung. Während Datenpunkte für bestimmte Indikatoren mit der Zeit an Gewichtung gewonnen haben, sank der Einfluss anderer Daten auf das Gesamtratingergebnis. Es zeigt sich, dass die Bewertungskriterien und die Schwerpunkte der Betrachtungen sich mit der Zeit anpassen und verändern. Dies liegt auch daran, dass sich die Rahmenbedingungen und zentralen Vorgaben nach wie vor verändern, etwa durch Beschlüsse wie die der COP21 in Paris im Jahre 2015.⁷² Diese Verschiebungen der Bewer-

⁶⁷ Vgl. Chatterji, A. et al. (2014), S. 6; Vgl. Billio, M. et al. (2021), S. 1431.

⁶⁸ Vgl. Berg, F./ Kölbel, J./ Rigobon R. (2022), S. 1328-1330.

⁶⁹ Vgl. Berg, F./ Kölbel, J./ Rigobon R. (2022), S. 1317.

⁷⁰ Vgl. Berg, F./ Kölbel, J./ Rigobon R. (2022), S. 1335.

⁷¹ Vgl. Mazzacurati, J. (2021), S. 107; Vgl. Scalet, S./ Kelly, T. F. (2010), S. 72; Vgl. Larcker, D. F. et al. (2022), S. 4; Vgl. Da Graca, A. (2022), S. 165.

⁷² Vgl. Escrig-Olmedo, E. et al. (2019), S. 10.

tungsschwerpunkte führen jedoch auch zu sich laufend ändernden ESG-Ratingnoten für die Unternehmen. Diese Veränderungen sind allerdings schwer von außen nachzuvollziehen und vergrößern die Unsicherheit bezüglich der tatsächlichen ESG-Leistung eines Unternehmens.

4.3 Muster bei ESG-Ratings

Neben den bereits genannten Gründen für Ratingabweichungen konnten in Studien über alle Ratingagenturen hinweg Muster identifiziert werden, die Auffälligkeiten bei Ratingergebnissen von verschiedenen Unternehmen zeigen.

Eine erste Erkenntnis ist, dass die Unternehmensgröße das Rating zu beeinflussen scheint und dass zwischen der Größe des Unternehmens und dem ESG-Rating eine Korrelation besteht.⁷³ Ergebnisse einer Untersuchung von *Da Graca* (2022) zeigen, dass ESG-Ratings tendenziell von der Größe der untersuchten Unternehmen abhängen und größere Unternehmen ein besseres Rating-Ergebnis erzielen als kleinere Unternehmen. In der Untersuchung wurden die Unternehmen anhand der Marktkapitalisierung in drei Größenklassen unterteilt. Ein Vergleich der Ratingergebnisse aller Größenklassen zueinander ergab vor allem für die Ratings von Refinitiv und ISS ESG beträchtliche Unterschiede und zeigt ein klares Muster bezüglich eines Größenzusammenhangs.⁷⁴

Dabei drängt sich die Frage auf, ob größere Unternehmen wirklich nachhaltiger sind als kleinere oder ob größere Unternehmen nur einen Vorteil aufgrund ihrer Struktur und der daraus entstehenden Möglichkeiten haben. Ein Erklärungsversuch lautet, dass größere Unternehmen mehr Ressourcen für die Erhebung und Veröffentlichung von ESG-Daten in Form einer aufwendigeren Nachhaltigkeitsberichterstattung haben und die größere Datenmenge die Ratingagenturen zu besseren Ratings verleitet.⁷⁵ Darüber hinaus stehen große Unternehmen häufiger im Fokus der Medien oder von NGO's und werden in deren Berichte mit aufgenommen, was ebenfalls zu einer größeren medialen Präsenz von Nachhaltigkeitsleistungen und Wahrnehmung in der Öffentlichkeit führt.⁷⁶

Dabei kann selbstverständlich nicht verallgemeinert werden, dass große Unternehmen unabhängig ihrer Möglichkeiten tatsächlich nachhaltiger sind als ihre kleineren Pendanten. Dennoch muss der Kritikpunkt angebracht werden, dass Medienberichte und Pressemeldungen für die ESG-Ratingerstellung herangezogen werden und Unternehmen bei nicht ausreichend genauer Prüfung der berichteten ESG-Leistung möglicherweise für die bessere PR-Strategie und nicht für die überlegene ESG-Leistung belohnt werden.⁷⁷

Neben der Größe des bewerteten Unternehmens scheint es auch Hinweise darauf zu geben, dass es eine Rolle spielt, wo das Unternehmen seinen Sitz hat. So zeigt sich, dass Unternehmen in Industrieländern im Durchschnitt bessere Ratingnoten erhalten als Unternehmen aus Emerging Markets.⁷⁸ Selbst innerhalb der Industrienationen gibt es Unterschiede, die sich daran zeigen, dass europäische Unternehmen durchschnittlich ein höheres ESG-Rating aufweisen als US-amerikanische.⁷⁹ Als mögliche Ursache wird auf die verschiedenen Rechtsvorschriften und abweichende Regulierung sowie die daraus resultierenden Anforderungen an die Berichterstattung verwiesen.⁸⁰

⁷³ Vgl. *Da Graca*, A. (2022), S. 167; Vgl. *Drempetic, S./ Klein, C./ Zwergel, B.* (2020), S. 335, Vgl. *Bams, D./ van der Kroft, B.* (2022), S. 7.

⁷⁴ Vgl. *Da Graca*, A. (2022), S. 165-166.

⁷⁵ Vgl. *Drempetic, S./ Klein, C./ Zwergel, B.* (2020), S. 335.

⁷⁶ Vgl. *Bams, D./ van der Kroft, B.* (2022), S. 7.

⁷⁷ Vgl. *Larcker, D. F. et al.* (2022), S. 1.

⁷⁸ Vgl. *Da Graca*, A. (2022), S. 165-167.

⁷⁹ Vgl. *Larcker, D. F. et al.* (2022), S. 5-6.

⁸⁰ Vgl. *Da Graca*, A. (2022), S. 166-167; Vgl. *Larcker, D. F. et al.* (2022), S. 5-6.

Neben dem Sitz des bewerteten Unternehmens scheint zudem auch die Herkunft der ESG-Ratingagentur Einfluss auf das Ratingergebnis zu haben. Wie *Chatterji et al. (2014)* herausfanden, legen US-amerikanische Ratingagenturen etwa mehr Wert auf die Dimension Soziales als europäische Ratingagenturen. Dabei kommt es auch innerhalb der Dimension Soziales zu deutlichen Unterschieden bei der Fragestellung, wie detailliert einzelne Teilaspekte betrachtet werden. Während einzelne Agenturen viele Fragen unter dem Begriff Beschäftigung zusammenfassen, unterscheiden andere stärker und berücksichtigen Einzelfaktoren wie Schulungsprogramme, Sicherheit der Mitarbeiter oder die Vielfalt innerhalb der Belegschaft.⁸¹

Gerade der zweite Punkt, die unterschiedlichen Herkünfte der Ratingagenturen sowie deren daraus resultierenden unterschiedlichen Bewertungen machen es derzeit wohl schwierig, die ESG-Ratings und den Ratingprozess zu vereinheitlichen. Hier kann es aber in Zukunft durch eine weitere Konsolidierung und Übernahmen über Landesgrenzen hinweg zu einer Anpassung innerhalb der Branche kommen. Zudem können internationale Vereinbarungen, etwa auf Ebene der *Vereinten Nationen* oder der *OECD*, die festgestellten Ratingdifferenzen zukünftig unter Umständen verringern.

Als drittes Muster kann genannt werden, dass es Branchen gibt, die generell besser bewertet werden als andere. Bei branchenbasierten ESG-Ratings fällt auf, dass Unternehmen, die ihre Haupttätigkeit in bestimmten Branchen haben, beispielsweise Banken und Telekommunikation, generell besser abschneiden als andere. Dabei ist nicht immer klar, ob die Unterschiede entstehen, weil es zwischen den Branchen einen generellen Qualitätsunterschied bei ESG-Faktoren gibt oder es an den angewandten ESG-Ratingmodellen, Inputparametern oder Gewichtungen im Ratingprozess liegt.⁸² Es ist aber offensichtlich, dass einige ESG-Ratingagenturen die Unternehmen einem Negativ-Screening unterziehen, was den Ausschluss derjenigen bedeutet, die bestimmte, als inakzeptabel definierte Sektoren oder Geschäftsfelder bedienen, die der Umwelt oder der Gesellschaft schaden.⁸³ Dies führt in der Untersuchung von *Escrig-Olmedo, Munoz-Torres und Fernandez-Izquierdo (2010)* dazu, dass ein Großteil der Agenturen die Beteiligung in der Tabakindustrie als wichtigstes Ausschlusskriterium festlegten, die Beteiligung am Uranabbau, Landminen oder der Embryonenforschung dagegen aber bei kaum einer Agentur von vorneherein zu einem Ausschluss führt.⁸⁴

5 Zusammenfassung und Ausblick

Mit Blick auf die zentrale Fragestellung dieser Untersuchung konnte herausgearbeitet werden, dass der Aussagegehalt von ESG-Ratings noch nicht das gewünschte Maß an Allgemeingültigkeit und Vergleichbarkeit bietet, den es für eine gezielte Steuerung der Finanzflüsse in nachhaltige Unternehmen und Geschäftsfelder benötigt.

Die undurchsichtige Raterstellung birgt aktuell die Gefahr, dass das Rating nicht auf die Nachhaltigkeitsaspekte abzielt, die den Investoren und Investorinnen wichtig sind. Jeder und jede muss vor einer Entscheidung individuell Nachhaltigkeitspräferenzen und Gewichtungen betrachten und bewerten⁸⁵, erst dann kann sich für eine Ratingagentur zur Betrachtung der ESG-Leistung entscheiden werden.

Die in diesem Beitrag aufgezeigten Divergenzen zwischen den ESG-Ratings bedeuten aber nicht, dass die Messung der ESG-Performance keine sinnvolle Aufgabe ist. Sie verdeutlichen aber, dass

⁸¹ Vgl. Chatterji, A. et al. (2014), S. 6.

⁸² Vgl. Larcker, D. F. et al. (2022), S. 5.

⁸³ Vgl. Billio, M. et al. (2021), S. 1429.

⁸⁴ Vgl. Escrig-Olmedo, E./ Muñoz-Torres, M. J./ Fernández-Izquierdo, M. A./ (2010), S. 449.

⁸⁵ Vgl. Schäfer, H. (2020), S. 382-383.

das Ergebnis durch viele Variablen beeinflusst wird. So ist die Auswahl der zugrundeliegenden Daten von wesentlicher Bedeutung. Auch die Methode, die eine Ratingagentur anwendet, sowie deren Definition von relevanten ESG-Risiken und die daraus folgende Gewichtung haben entscheidenden Einfluss auf das Rating-Endergebnis.

Generell besteht derzeit noch ein deutlicher Bedarf an Standardisierung, angefangen mit einer einheitlichen ESG-Definition. Hier werden auch die Regulierungsbehörden gefragt sein, die die Regeln und Methoden der Nachhaltigkeitsberichterstattung weiter vereinheitlichen müssen, vergleichbar mit den Regelungen für die Rechnungslegungsstandards.⁸⁶ Zudem kann erwartet werden, dass es in Zukunft zu einer weiteren Konsolidierung des Marktes durch Fusionen und Übernahmen kommen wird. Auch deshalb wird es in Zukunft wichtig sein, die Transparenz im Bewertungsprozess weiter zu fördern, Bewertungsskalen und Bewertungskriterien zu vereinheitlichen und allgemeingültige Risikodefinitionen zu erstellen. Dies wird die Akzeptanz und Glaubwürdigkeit gegenüber privaten und institutionellen Investoren in Zukunft weiter erhöhen.

Abschließend kann mit diesem Beitrag festgestellt werden, dass die ESG-Ratingnoten aktuell noch nicht die gewünschte Vereinfachung darstellen, wenn es darum geht, die ESG-Leistung eines Unternehmens zu beurteilen. Die bestehenden Informationsasymmetrien werden zwar ein Stück weit beseitigt, jedoch ist nach wie vor viel Recherchearbeit notwendig, wenn aus der Vielzahl an ESG-Bewertungsergebnissen eine verlässliche Nachhaltigkeitseinschätzung zu einem Unternehmen erarbeitet werden soll. Hier ist nun die weitere Entwicklung zu beobachten und zukünftig zu prüfen, ob eine weitere Standardisierung durch zentrale Vorgaben stattfindet oder sich die ESG-Ratingergebnisse in Zukunft durch eine Marktkonsolidierung und stärkeren Druck von Seiten der Anleger und Anlegerinnen einander angleichen.

⁸⁶ Vgl. Billio, M. et al. (2021), S. 1445.

Literaturverzeichnis

- Amel-Zadeh, A./ Serafeim, G. (2018): Why and How Investors Use ESG Information: Evidence from a Global Survey. In: *Financial Analysts Journal*, Vol. 74 (3), pp. 87-103.
- Avetisyan, E./ Hockerts, K. (2017): Consolidation of the ESG Rating Industry as enactment of institutional retrogression. In: *Business Strategy and the Environment*, Vol. 26 (3), pp. 316-330.
- Bams, D./ van der Kroft, B. (2022): Tilting the Wrong Firms? How Inflated ESG Ratings Negate Socially Responsible Investing Under Information Asymmetries. In: MIT Center for Real Estate Research Paper No. 22/12, S. 1-25.
- Berg, F./ Heeb, F./ Kölbel, J. (2022): The Economic Impact of ESG Ratings. Zuletzt überarbeitet 2023; URL: <https://ssrn.com/abstract=4088545> (Stand 10.04.2023).
- Berg, F./ Kölbel, J./ Rigobon R. (2022): Aggregate Confusion: The Divergence of ESG Ratings. In: *Review of Finance*, Vol. 26, No. 6, pp. 1315-1344.
- Billio, M./ Costola, M./ Hristova, I./ Latino, C./ Pelizzon, L. (2021): Inside the ESG ratings: (Dis)agreement and performance. In: *Corporate Social Responsibility and Environmental Management*, Vol. 28, No. 5, pp. 1426-1445.
- Bundesanstalt für Finanzdienstleistungsaufsicht (Hrsg.) (2020): Merkblatt im Umgang mit Nachhaltigkeitsrisiken. URL: https://www.bafin.de/SharedDocs/Downloads/DE/Merkblatt/dl_mb_Nachhaltigkeitsrisiken.pdf?__blob=publicationFile&v=14 (Stand: 02.04.2023).
- Chatterji, A./ Levine, D. (2006): Breaking down the Wall of Codes: Evaluating Non-Financial Performance Measurement. In: *California Management Review*, Vol. 48, No. 2, pp. 29-51.
- Chatterji, A./ Durand, R. / Levine, D./ Touboul, S. (2014): Do Ratings of Firms Converge? Implications for Strategy Research. In: IRLE - Institute for Research on Labor and Employment (Hrsg.), IRLE Working Paper No. 107-14, 2014.
- Christensen, D./ Serafeim, G./ Sikochi, A. (2022): Why is Corporate Virtue in the Eye of The Beholder? The Case of ESG Ratings. In: *The Accounting Review*, Vol. 97, No. 1, pp. 147-175.
- Da Graca, A. (2022): Unterschiede bei ESG-Ratings in der Praxis. In: *CORPORATE FINANCE*, Nr. 05-06, 2022, S. 158-167.
- Drempetic, S./ Klein, C./ Zwergel, B. (2020): The Influence of Firm Size on the ESG Score: Corporate Sustainability Ratings Under Review. In: *Journal of Business Ethics*, Band 167 / Ausgabe 2, 2020, S. 333-360.
- Escrig-Olmedo, E./ Fernández-Izquierdo, M. A./ Ferrero-Ferrero, I./ Rivera-Lirio, J. M./ Muñoz-Torres, M. J. (2019): Rating the Raters: Evaluating how ESG Rating Agencies Integrate Sustainability Principles. In: *Sustainability*, Vol. 11, No. 3, Article 915.
- Escrig-Olmedo, E./ Muñoz-Torres, M. J./ Fernández-Izquierdo, M. A./ (2010): Socially responsible investing: sustainability indices, ESG rating and information provider agencies. In: *International Journal of Sustainable Economy*, Vol. 2, No. 4, pp. 442-461.
- ESMA (Hrsg.) (2021): ESMA Report on Trends, Risks and Vulnerabilities No 1, 2021, S. 105-115. URL: https://www.esma.europa.eu/sites/default/files/library/esma50-165-1524_trv_1_2021.pdf (Stand: 01.04.2023).
- Horn, M. (2023): The Influence of ESG Ratings on Idiosyncratic Stock Risk: The Unrated, the Good, the Bad, and the Sinners. In: *Schmalenbach Journal of Business Research* (2023). <https://doi.org/10.1007/s41471-023-00155-1>.
- Larcker, D. F./ Pomorski, L./ Tayan, B./ Watts, E. M. (2022): ESG Ratings: A Compass without Direction. In: Stanford Closer Look Series. Corporate Governance Research Initiative (2022).
- Leinweber, L./ Kögler, A. (2023): Was taugen ESG-Ratings wirklich? In: *FINANCE - Der Markt für Unternehmen und Finanzen*, Heft März/April 2023, S. 64-65.

- Leker, J./ Frieden, F. (2022): ESG – Diffuse Kennzahlen als Entscheidungsgrundlage? In: CORPORA-FINANCE, Nr 11-12, 2022, S. M1.
- Lykkesfeldt, P./ Kjaergaard, L. L. (2022): Investor Relations and ESG Reporting in a Regulatory Perspective - A Practical Guide for Financial Market Participants, Cham, Schweiz.
- Mazzacurati, J. (2021): ESG ratings: Status and key issues ahead. In: ESMA (Hrsg.), Report on Trends, Risks and Vulnerabilities, No. 1, 2021, S. 105-115.
- MSCI (Hrsg.) (2020): MSCI ESG Ratings. URL: <https://www.msci.com/documents/1296102/21901542/MSCI+ESG+Ratings+Brochure-cbr-en.pdf> (Stand: 04.05.2023).
- MSCI (Hrsg.) (2024): ESG Ratings & Climate Search Tool. URL: <https://www.msci.com/our-solutions/esg-investing/esg-ratings-climate-search-tool> (Stand 11.02.2024).
- Ossenbrink, N. (2019): Nachhaltigkeitsratings als Ergänzung der traditionellen Finanzanalyse. In: Everling, O./ Lampe, J. (Hrsg.), Rating von Vermögensverwaltungen, Wiesbaden, S. 129-143.
- Refinitiv (Hrsg.) (2022): ENVIRONMENTAL, SOCIAL AND GOVERNANCE SCORES FROM REFINITIV. URL: https://www.refinitiv.com/content/dam/marketing/en_us/documents/methodology/refinitiv-esg-scores-methodology.pdf (Stand 05.05.2023).
- Refinitiv (Hrsg.) (2024): LSEG ESG Scores. URL: <https://www.lseg.com/en/data-analytics/sustainable-finance/esg-scores> (Stand 11.02.2024).
- S&P Global (Hrsg.) (2023a): ESG Scores. URL: <https://www.spglobal.com/esg/solutions/data-intelligence-esg-scores?> (Stand: 04.05.2023).
- S&P Global (Hrsg.) (2023b): ESG Evaluation. URL: https://www.spglobal.com/_assets/documents/ratings/esg/esg_evaluation_brochure_digital.pdf (Stand: 04.05.2023).
- S&P Global (Hrsg.) (2024): ESG Scores. URL: <https://www.spglobal.com/esg/solutions/data-intelligence-esg-scores> (Stand 11.02.2024).
- Scalet, S./ Kelly, T. F. (2010): CSR Rating Agencies: What is Their Global Impact? In: Journal of Business Ethics, Band 94 / Ausgabe 1, 2010, S. 69-88.
- Schäfer, H. (2020): 20. Green Economy, Green Deal und Sustainable Finance – Die zentrale Rolle von Nachhaltigkeitsratings. In: Everling, O. (Hrsg.), Social Credit Rating, Wiesbaden, S. 367-389.
- Sustainalytics (Hrsg.) (2023): ESG Risk Ratings – A consistent approach to assess material ESG risk. URL: <https://www.sustainalytics.com/esg-data> (Stand: 04.05.2023).
- Sustainalytics (Hrsg.) (2024): Company ESG Risk Ratings. URL: <https://www.sustainalytics.com/esg-ratings> (Stand 11.02.2024).
- Tschirschwitz, D./ Klemstein, F./ Stein, B./ Rodehorst, V. (2022): A Dataset for Analysing Complex Document Layouts in the Digital Humanities and Its Evaluation with Krippendorff's Alpha. In: Andres, B./ Bernard, F./ Cremers, D./ Frintrop, S./ Goldlücke, B./ Ihrke, I. (Hrsg.), Pattern Recognition - 44th DAGM German Conference, DAGM GCPR 2022, Konstanz, Germany, September 27–30, 2022, Proceedings, Cham, Schweiz, S. 354-374.
- Vereinte Nationen (Hrsg.) (2004): Who Cares Wins – Connecting Financial Markets to a Changing World. URL: https://www.unepfi.org/fileadmin/events/2004/stocks/who_cares_wins_global_compact_2004.pdf (Stand: 02.04.2023).
- Windolph, S. E. (2011): Assessing Corporate Sustainability Through Ratings: Challenges and Their Causes. In: Journal of Environmental Sustainability: Vol. 1, No. 1, pp. 61-80.
- Wohnhaas, B. (2023): ESG-Rating im Fokus. In: FINANCE - Der Markt für Unternehmen und Finanzen, Sonderbeilage November/Dezember, 2021, S. 26.

Hochschule für Finanzwirtschaft & Management
Simrockstraße 4
53113 Bonn

