

Secondary Publication



Oehler, Andreas; Neuss, Charlotte

ESG disclosure vs. ESG ratings : Consistent information value?

Date of secondary publication: 24.09.2025

Version of Record (Published Version), Article

Persistent identifier: urn:nbn:de:bvb:473-irb-110449x

Primary publication

Oehler, Andreas; Neuss, Charlotte (2025): ESG disclosure vs. ESG ratings : Consistent information value?, in: International review of financial analysis, Amsterdam [u.a.]: Elsevier, Vol. 107, Nr. 104623, pp. 1–24, doi: 10.1016/j.irfa.2025.104623.

Legal Notice

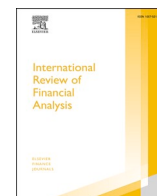
This work is protected by copyright and/or the indication of a licence. You are free to use this work in any way permitted by the copyright and/or the licence that applies to your usage. For other uses, you must obtain permission from the rights-holders.

This document is made available under a Creative Commons license.



The license information is available online:

<https://creativecommons.org/licenses/by/4.0/legalcode>



ESG disclosure vs. ESG ratings: Consistent information value?[☆]

Andreas Oehler^{*}, Charlotte Neuss

Department of Finance, Bamberg University, Bamberg, Germany

ARTICLE INFO

JEL codes:

G10
G11
G12
G24
D83

Keywords:

ESG rating
Sustainability disclosure
ESG reporting
Information value
Corporate governance
CSR reporting

ABSTRACT

We perform textual analysis of annual reports and examine the consistency between firms' ESG rating and firms' ESG disclosure. The paper contributes to the ESG literature by combining two mostly separately investigated research topics: the ESG ratings and the ESG disclosure. This is important as market participants use both sources for decision-making. We develop two different word lists to analyze the ESG information value from a rating agency-perspective and from the academic literature perspective. Our results show that using a comprehensive ESG-related word list from academic studies can explain a substantial proportion of several ESG ratings, indicating that market participants can use firm disclosure as an approximation for several ESG ratings. Given that ESG rating agencies use public information, our study sheds a light on the information value from different information channels which is important for market participants. Moreover, the differences in the information value reveal the difficulties in using ESG ratings and the issue of rating dispersions.

1. Introduction

Environmental, Social, and Governance (ESG) issues have been one of the fastest-growing trends globally and particularly in the financial industry. The worldwide development in adopting ESG principles is driven by regulatory incentives and increased demand for sustainable companies and their ESG disclosures. ESG ratings provide an aggregated assessment of firms' ESG activities and serve as a crucial information source for investors (e.g., Bialkowski & Starks, 2016; Hartzmark & Sussman, 2019; Horn, 2023; Horn & Oehler, 2024). These ratings incorporate numerous ESG indicators; however, there are significant differences among ESG rating providers¹ (e.g., Berg, Kölbl, & Rigobon, 2022; Chatterji et al., 2016; Clementino & Perkins, 2021; Horn & Oehler, 2024; Oehler & Horn, 2025). This variability is particularly noteworthy because ESG ratings are derived from publicly available

information, such as firms' self-reported ESG disclosure, media news, and other information about the firm (Billio et al., 2021).² ESG ratings are costly but provide a professional ESG assessment from ESG rating analysts.

Another information channel is firms' ESG disclosures. Empirical evidence suggests that investors find firms' ESG disclosures relevant and useful (e.g., Amel-Zadeh & Serafeim, 2018; Papoutsis & Sodhi, 2020; Verbeeten et al., 2016). The disclosure with its three ESG pillars is helpful in assessing firms' current and future ESG activities, opportunities, and threats because ESG topics are closely related to firms' regular business activities and strategic alignment (Albarrak et al., 2019; Dhaliwal et al., 2011, 2012). New ESG disclosure is followed by a significant investor reaction (Aureli et al., 2020). Companies' reports are free of charge and downloadable from their websites. Nevertheless, firms have incentives to present a more favorable picture of their

[☆] We want to thank Matthias Horn, Stefan Wendt and anonymous referees for the valuable comments and suggestions. In addition, we want to thank the participants at the Behavioural Finance Working Group 2024 Conference, the CINSC 3rd Conference on International Finance, Sustainable and Climate Finance and Growth, the European Financial Management Association 2024 Annual Meeting, and the Financial Reporting and Business Communication Conference 2024 for helpful comments and suggestions. All remaining errors are our own.

^{*} Corresponding author at: Bamberg University, Kaerntenstrasse 7, 96052 Bamberg, Germany.

E-mail address: andreas.oehler@uni-bamberg.de (A. Oehler).

¹ Berg et al. (2022) focus on the divergence between ESG scores of different rating agencies and argue that the differences can mainly be explained by measurement (56 %), i.e., in the process of measuring an attribute, different indicators are used (see also Abhayawansa & Tyagi, 2021). The second reason is scope (38 %), i.e., a different set of attributes is considered. Dimson et al. (2020) find that data providers assign different weights of underlying pillar information.

² We use the term ESG disclosure instead of sustainability disclosure. Although there is some overlap between both concepts, our analysis is not restricted to sustainability but also includes content from the G-pillar.

business.

Compared to ESG disclosures, ESG ratings may be more convenient, easier to use in assessment tools, and timesaving. Additionally, ESG ratings can cover critical information, such as the identification of greenwashing and the strategic positioning within the industry. Since ESG rating agencies are not transparent about their metrics, market participants must rely on the quality of the ESG ratings without fully understanding the included factors. Additionally, market participants face challenges due to the lack of standardized ESG disclosures and the varying information value of ESG ratings. One reason for this inconsistency is the different regulatory requirements around the world (e.g., Baldini et al., 2018; Buallay, 2019; Oehler & Horn, 2025; Serafeim & Yoon, 2022; van der Laan Smith et al., 2005). Despite the lack of transparency, market participants use the ESG disclosures and ESG ratings for decision-making (e.g., Galema & Gerritsen, 2025; Hartzmark & Sussman, 2019; Verbeeten et al., 2016).

While there is a broad literature covering ESG ratings as well as ESG disclosure, the comparison of these two information sources remains limited. A positive relationship between firm disclosure and sustainability scores is generally established (see e.g., Ferjančić et al., 2024; Huang et al., 2024; Kartal et al., 2024; Khatri & Kjærland, 2023; Papoutsis & Sodhi, 2020; Schimanski et al., 2024). Moreover, ESG disclosure can reduce rating disagreement (Huang et al., 2024; Kimbrough et al., 2024; Liu, 2022), which underlines the importance of ESG disclosures. Understanding the relationship between the two information channels – ESG disclosure and ESG ratings – is particularly important for market participants and regulatory authorities, especially given the criticism surrounding ESG rating agencies. However, due to reliance on disclosure scores, the consistency across different information channels remains partly unexplored. The black-box nature of ESG ratings and other sustainability scores means that users often cannot discern which information is incorporated into these measures or how much value ESG rating agencies add beyond publicly available information. Therefore, it is crucial to examine the extent to which ESG ratings reflect firms' ESG disclosures to enhance transparency in two key information channels (see Kartal et al., 2024). This requires overcoming the black-box problem of disclosure scores by data providers.

Our study focuses on annual reports as the primary disclosure channel, considers multiple ESG ratings, and includes a textual analysis that offers transparency about the input. This allows important conclusions for researchers, regulators, and different stakeholder groups and reduces the dependency on data providers of disclosure scores. We extend the literature by comparing the ESG disclosure and the (following) ESG rating for a dataset of German listed firms. Germany was chosen due to its reporting regulation that follows the Directive 2014/95/EU since 2017.³ Additionally, Germany shows stakeholder orientation and is characterized by strong unions, institutionalized collective bargaining systems, extensive provision of social welfare and employment protection. Our study contributes to the literature and practice by providing insights into the consistency between ESG ratings and firm's ESG disclosure from annual reports.

Given the critique of available ESG scores,⁴ we create our own ESG measures based on firm disclosure and distinguish between two word lists. We follow the *MSCI KLD ESG Rating Guideline* (hereinafter, MSCI

word list) and extract keywords. The guideline outlines the ESG categories considered in the rating process of MSCI KLD. We assume that using this MSCI word list on our sample of annual reports will highly explain firms' ESG ratings from MSCI KLD and represents the rating agency approach. Second, we use ESG keywords from several studies to create a comprehensive ESG-related word list from the academic perspective (hereinafter, academic word list). This word list incorporates the results from twelve studies⁵ from 1982 to 2022 as well as the Sustainable Development Goals (SDG).⁶ Moreover, we distinguish between two textual variables: (1) the absolute word count, which includes repetitions, and (2) the presence of keywords, represented as dummy variables for each keyword.

Our results indicate that the information from the annual report can serve as an approximation of several ESG ratings. Market participants who rely on ESG ratings for decision-making are limited to rated firms, potentially leading to selection bias. If these market participants can extract substantial ESG information from annual reports on their own, this reduces their dependence on ESG rating agencies and increases their flexibility in decision-making. We show that ESG disclosure can explain a substantial part of several ESG ratings. Comparing both word lists shows that the academic word list has a higher explanatory power for several ratings. Nevertheless, ESG ratings face criticism for various reasons. Our findings are important because ESG ratings are used for investment decisions in practice.

Although we assumed that the MSCI word list would explain a higher portion of firms' ESG ratings from MSCI KLD compared to ratings from the other rating agencies, the results do not support this assumption. Surprisingly, the explained variance of the MSCI KLD ESG rating is lower than that of other ESG ratings, even though the disclosure variable was constructed based on MSCI's own rating framework. This outcome suggests several possibilities. First, it indicates that the MSCI word list, while aligned with MSCI's stated rating categories, may only marginally enhance the transparency of how ratings are derived. Second, it could imply that MSCI KLD incorporates additional layers beyond what is reflected in public disclosures. This implies that MSCI KLD may weigh qualitative factors or internal assessments more heavily, weakening the direct influence of public disclosures on the rating outcome. Furthermore, the stronger explanatory power of the MSCI disclosure variable for other ESG ratings may point to greater alignment between MSCI's disclosure framework and the data-gathering processes of other rating agencies. This could reflect methodological convergence across rating agencies or differences in how closely each rating system relies on textual disclosures.

When comparing the textual variables, we find that the MSCI word list explains a higher proportion of several ESG ratings when using the presence of keywords instead of the absolute word count. In contrast, for the academic word list, we find the opposite results, i.e., a higher proportion of the ratings is explained when using the absolute count of keywords rather than their presence. This contrast provides insight into how different types of information may be processed during the rating creation process and suggests that the MSCI word list may function more like a checklist.

The results highlight that annual reports can serve as useful proxies for specific ESG ratings when applying a sophisticated ESG-related word list. Moreover, the observed differences in information value reveal the

³ Since 2017, selected German firms have been obliged to report on ESG, although the specific content is not defined. Selected firms are defined as firms that are oriented towards the capital market, that have more than 500 employees on average in the course of a year (§289b German commercial code) and have more than €20 million in total assets or more than €40 million in sales. All firms in our sample have to follow the directive 2014–95-EU due to their capital market orientation.

⁴ The critique mostly relates to the divergence in ESG scores of different data providers, the methodology, missing data, and unannounced changes (see, e.g., Benuzzi et al., 2024; Oehler & Horn, 2025; Sahin et al., 2022).

⁵ Aldridge and Martin (2022); Baier et al. (2020); Borms et al. (2021); Bouten et al. (2011); Clarkson et al. (2008); Cormier and Magnan (2003); Gamerschlag et al. (2011); Kouloukoui et al. (2019); Papoutsis and Sodhi (2020); Tagesson et al. (2009); Wiseman (1982).

⁶ The Sustainable Development Goals (SDG) reflect a framework of 17 main goals, 169 targets together with indicators. In 2015, the 2030 Agenda for Sustainable Development started and is adopted by all United Nations Member States (see e.g., <https://sdgs.un.org/goals>; last access September 28, 2023, SDG (2022)).

challenges in using ESG ratings and the issue of rating dispersion.

This study is part of the limited area dealing with ESG ratings as the dependent variable. We investigate the main drivers for the relationship between the ESG rating and the ESG disclosure. Although the weight of the pillars is different between ESG rating providers, the G-pillar always shows the lowest contribution to the rating. We further analyze sub-categories in each pillar (e.g., ‘pollution and waste’ in the E-pillar) and the results show higher explanatory power for ‘climate change’, ‘product liabilities’, ‘human capital’, and ‘corporate governance’ (although there are differences between both word lists). This finding suggests that certain ESG topics may have greater weight in the rating process. The high explanatory power of ‘climate change’, for example, indicates that firms’ disclosures on climate-related issues are particularly influential for ESG ratings. This could reflect both the growing importance of climate risks in investor decision-making and the emphasis placed on climate transparency by rating agencies.

The paper is structured as follows. In the next section, we review the literature and introduce ESG measures. We describe our data and methodological approach in section 3. In section 4, we present our results, provide robustness checks, and discuss the results. Section 5 concludes.

2. Literature review: ESG disclosure and ESG ratings

Different theories explain the relationship between sustainable practices of firms and the role of different stakeholders. Transparent disclosure of (ESG) information can enhance a firm’s reputation, provide access to capital at lower costs, and improve competitive advantages (e.g., Ben-Porath et al., 2018; Gillan et al., 2021). According to voluntary disclosure theory, firms with better environmental performance seek to produce credible direct disclosures, i.e., disclosure with a focus on objective measures that are hard to mimic by less environmental conscientious firms (Verrecchia, 1983). Agency theory suggests that greater disclosure reduces information asymmetries. Another theoretical strand is the socio-political theory including legitimacy theory. ESG disclosure can generate competitive advantages, influence reputation and the corporate image, and satisfy the information needs of investors and other stakeholders (Tamimi & Sebastianelli, 2017). Following legitimacy and stakeholder theory, firms use disclosure as a legalized tool (e.g., Haniiffa & Cooke, 2005). In this context, firms disclose more information due to pressure from stakeholders.

In practice, ESG disclosure has several standards and different voluntary guidelines or frameworks that want to ensure sophisticated reporting, for example, the Global Reporting Initiative (GRI), which is recommended explicitly by German regulation, or the International Sustainability Standards Board (ISSB). However, firms can decide the level of detail in their disclosure, as there is no mandatory specification for the exact content. To comply with the Directive 2014/95/EU, firms typically integrate disclosures related to their societal and environmental impact into their annual reports; alternatively, firms provide a separate report. Auditors currently only verify the existence of the ESG report, not its content. German firms follow the IFRS, but the IFRS S1 and S2 requirements will begin to apply in 2024. Generally, firms are expected to report on key topics covering environmental practices, social information, and corporate governance systems. Environmental practices typically include measures of emissions, waste, pollution, water and energy consumption, renewable energy production, climate change risks, and environmental conservation. Social information covers managing relationships with stakeholders, including labor relations, product liability, supply chain management, community investment, labor and human right policies, and health and safety policies. Governance disclosure includes board structure, auditing procedures, ethical principles, and shareholder rights (e.g., Raimo et al., 2021). Governance criteria also include standards to ensure transparent accounting models that promote integrity and diversity (Benuzzi et al., 2024). Reading ESG disclosures requires little experience, but the broad

range of topics and varying formats make it difficult to compare firms (e.g., Amel-Zadeh & Serafeim, 2018). Moreover, the long-term prospects are hard to quantify and may be intangible (Christensen et al., 2021).

Different data providers offer disclosure scores based on firm reporting, for example the London Stock Exchange Group (LSEG) (formerly Refinitiv and Thomson Reuters) ASSET4 ESG database⁷ or Bloomberg’s ESG disclosure score.⁸ However, the availability of ESG scores is limited and often contains missing information, complicating the analysis and potentially leading to selection bias in ESG-related studies. Some scholars therefore question the reliability of ESG scores (e.g., Berg, Kölbl, & Rigobon, 2022; Dremptic et al., 2020; Sahin et al., 2022). Dremptic et al. (2020) find that more available data, regardless of whether it is positive or negative, raises the firm’s overall sustainability assessment from Thomson Reuters. Furthermore, managers may take strategic actions to improve their ESG score. This damages the long-term value of their companies and underlines the critique of ESG scores (Avetisyan & Hockerts, 2017).

To address the limitations of using ESG scores from data providers, textual analysis offers a way to measure ESG disclosure by analyzing the frequency of specific words or sentences in documents (e.g. word lists like in Loughran & McDonald, 2011; Tetlok, 2007). Researchers derive word lists for content analysis based on frameworks like the GRI (e.g., Lokuwaduge & Heenetigala, 2017; Verbeeten et al., 2016) or create individual word lists. The GRI guideline, published in 1999 and revised multiple times, covers 11 principles (Transparency, Inclusiveness, Auditability, Completeness, Relevance, Sustainability Context, Accuracy, Neutrality, Comparability, Clarity, and Timeliness), to ensure a comprehensive reporting of economic, environmental, and social activities, facilitating the comparison across firms and over time. The GRI is one of the most widely used ESG reporting guidelines.⁹

Manual content analysis based on the GRI has been used in various studies. Bouten et al. (2011) investigate the comprehensiveness of sustainability reporting of Belgian firms and find a low level of comprehensiveness. Tagesson et al. (2009) use company websites in addition to annual reports and find a positive correlation of firm size and profitability with the content on corporate websites. Schadewitz and Niskala (2010) examine the value relevance of disclosure and find that the use of GRI is an explanatory factor for Finnish firms’ market value. Gameschlag et al. (2011) develop a list of 32 keywords based on GRI to examine the determinants of voluntary disclosure. Plumlee et al. (2015) study the relationship between environmental disclosure quality (based on the GRI) and firm value, finding evidence for a positive association. Kouloukoui et al. (2019) focus on climate risk disclosure based on the GRI.

Other studies use individually created word lists. Wiseman (1982) designs an environmental disclosure index with four categories and 18 items, assigning scores based on whether the disclosure is quantitative or qualitative.¹⁰ A similar approach is conducted by Cormier and Magnan (2003) using 39 items from six categories to examine the determinants of environmental disclosure. Loughran et al. (2009) find that firms using ethics-related terms are more likely to score poorly on corporate governance measures. Clarkson et al. (2013) use the word list from Clarkson et al. (2008) with 95 items to examine voluntary environmental disclosure. The authors divide their items into “strong” and

⁷ The database is used in several studies (e.g., Ding et al., 2022; Dremptic et al., 2020; Dyck et al., 2019; Khatri & Kjærland, 2023; Mbanye & Muchenje, 2022).

⁸ The Bloomberg ESG score is used for example in the studies from Manita et al. (2018), Rajesh and Rajendran (2020), and Raimo et al. (2021).

⁹ Further guidelines are Sustainability Accounting Standards Board, Carbon Disclosure Project, International Integrated Reporting Council, and Carbon Disclosure Standards Board (Threlfall et al., 2020).

¹⁰ Point assignment: three for quantitative, two for non-quantitative, one for general terms, and no points for no disclosure.

“soft” and find value-relevant information about the historical environmental performance of U.S. companies. A combination of different guidelines is used by [Reverte \(2016\)](#) who uses data from OCSR that incorporates several norms beside the GRI. [Baier et al. \(2020\)](#) state that a useful list of words covering all ESG topics cannot be found in the literature so far and offer a word list with 482 items to quantify ESG reporting. Their results indicate a low share of sustainability-related words on total words. [Heichl and Hirsch \(2023\)](#) expand the word list from [Baier et al. \(2020\)](#) and investigate how precisely ESG topics are disclosed in sustainability reports. The authors create a sustainable fingerprint by using ESG and time reference dictionaries. Using artificial intelligence techniques, [Aldridge and Martin \(2022\)](#) find strong statistical significance of ESG terms predictability of future return characteristics. [Borms et al. \(2021\)](#) examines news about ESG performance using semi-supervised text mining, finding no performance loss when applying news-based sustainability screening.

ESG ratings allow market participants to screen firms ESG activities easily through an aggregated assessment. The theoretical foundation for (ESG) rating agencies is provided by new institutional economics, which emphasizes the importance of institutions in reducing information asymmetries by acting as intermediaries and reducing transaction costs ([Eggertsson, 2013](#)). ESG ratings are widely used in academic research (e. g., [Bird et al., 2007](#); [Galema & Gerritsen, 2025](#); [Horn, 2023](#); [Horn & Oehler, 2024](#); [Hull & Rothenberg, 2008](#); [Kimbrough et al., 2024](#); [Lins et al., 2017](#); [Mishra & Modi, 2013](#); [Muck & Schmidl, 2024](#); [Oehler & Horn, 2025](#); [Pelster et al., 2024](#)). However, there is no clear definition of the ESG assessment, and the values used by rating agencies are not open to the public ([Abhayawansa & Tyagi, 2021](#)). The weights and the data sources are considered intellectual property of a rating agency. Investors must inform themselves about the structure and information, but it is not transparent what the ESG ratings measure and how ESG rating agencies are creating those ratings. [Drempetic et al. \(2020\)](#) find that firm size variables highly influence ESG scores from ASSET4 ESG database, suggesting the ESG scores depend on the amount of available information, which is generally higher for larger firms, rather than a realistic picture of sustainability actions. ESG rating agencies assess companies based on several principles grouped into different subcategories. Although it can be assumed that rating agencies use similar methods, there are differences in the rating measurement, e.g., the topics included in the subcategories and their weights. [Guerrero and Viteri \(2025\)](#) examine the main components of ESG scores from LSEG and find that the scores are mainly driven by indicators measuring policies, processes, and initiatives. Moreover, the evaluation of a firm is based on its exposure to risk and its capability to manage the risks ([Berg, Heeb, & Kölbel, 2022](#)). Raw ESG scores are weighted and benchmarked against industry peers (e.g., [Tarnaud & Zakriya, 2023](#)). The ESG ratings can change, if firms’ ESG practices and specific issues improve or deteriorate, or if rating analysts adjust the industry peer group ([Berg, Heeb, & Kölbel, 2022](#)). These changes are followed by (delayed) stock price adjustments, underlying that ESG ratings are considered relevant by investors ([Galema & Gerritsen, 2025](#); [Glück et al., 2022](#)).

Combining the fields of ESG rating and ESG disclosure, [Christensen et al. \(2022\)](#) focus on the extent of firms’ ESG disclosure and the (dis)agreement among different ESG ratings. Unlike earnings forecasts in equity markets, their study finds that greater disclosure leads to greater rating divergence. [Kimbrough et al. \(2024\)](#) find that disagreement among ESG rating agencies is lower for firms that voluntarily disclose ESG information. In contrast, their study shows that longer reports are associated with reduced disagreement among the ESG rating agencies, while a more positive tone relates to higher disagreement. Dividing the ESG disclosure by its content, [Liu \(2022\)](#) finds that quantitative ESG disclosure of Chinese firms reduces rating disagreement among Chinese rating agencies, with environmental and social issues contributing more to disagreement than governance disclosures. [Huang et al. \(2024\)](#) explain that sufficiently specific reporting regulations reduce ESG rating disagreement by improving ESG disclosure. [Eng et al. \(2022\)](#) compare

sustainability disclosures from Sustainability Accounting Standards Board (SASB) Navigator and disclosure scores from Bloomberg finding that both sources provide incremental information and suggesting that it would be beneficial to harmonize standards for sustainability disclosures. [Ferjančić et al. \(2024\)](#) identify ESG topics and examine their evolution over time showing the influence of reporting regulation on ESG disclosure. Moreover, the authors find a positive correlation between the ESG topics and ESG scores from LSEG. [Papoutsis and Sodhi \(2020\)](#) create sustainability constructs from sustainability reports and significantly explain measures of sustainability performance from Bloomberg and Dow Jones Sustainability Index. [Khatri and Kjørland \(2023\)](#) show that issuing sustainability reports and adhering to the GRI standard is positively related to Refinitiv’s environmental pillar score. [Kartal et al. \(2024\)](#) apply a novel super learner algorithm and demonstrate that ESG principles have strong predictive power of ESG scores. These principles are derived from ESG disclosure in Türkiye in 2022, which provides a unique research setting due to recent changes in disclosure regulations. The authors call for further research in other countries, as structured ESG reports are currently only available in Türkiye.

These empirical studies highlight the importance of examining the relationship between firm disclosure and ESG ratings, as well as sustainability performance measures. Based on the relevance of ESG ratings for financial decision-making, we aim to answer the research question, how well ESG disclosure can explain ESG ratings. This investigation is important, as it contributes to a clearer understanding of ESG ratings. Additionally, we aim to address the “black box” problem of ESG disclosure scores by applying textual analysis, which enhances transparency and reduces reliance on third-party data providers. [Schimanski et al. \(2024\)](#) examine the relationship between ESG communication and ESG ratings and find that increasing ESG communication leads to higher ESG ratings. We extend their study by incorporating data from the largest ESG rating agencies, thereby offering a broader perspective. Our study further differentiates itself by comparing two perspectives, i.e., academic vs. rating agency perspective, allowing us to differentiate between varying amounts and content of ESG disclosure. Using a comprehensive word list grounded in prior literature, we categorize results by ESG topics and pillars across an 18-year period.

Following the positive relationship between ESG disclosure and ESG ratings demonstrated in other studies, we anticipate that the disclosure variables will exhibit substantial explanatory power. We assume this will be particularly pronounced within the subset of firms rated by MSCI KLD. This is the case since one of our word lists is created based on the MSCI KLD ESG rating guideline, i.e., the selection of topics included in the rating is considered within the word list. We argue that the greater the consistency between the disclosure and the ratings, the less value the ESG rating agencies add. If the information channels show high consistency, market participants can use firm reporting as an approximation for ESG ratings and the added value provided by rating agencies is relatively small.

3. Data and methodology

3.1. ESG ratings and stock data

As previous studies have shown, country-specific characteristics influence ESG adoption ([Liang & Renneboog, 2017](#)) and ESG disclosure ([Heichl & Hirsch, 2023](#); [van der Laan Smith et al., 2005](#)). ESG activities are often based on a firm’s geographical location (e.g., [Gillan et al., 2021](#)). We focus on a single country to ensure a consistent political and societal background. Our sample covers German stocks listed on the CDAX between 2002 and 2019, which comprises all prime and general standard equities in Germany. As we conduct textual analysis of annual reports, we are limited by the availability of these reports. We start from 2002 with a lower number of reports available, and include all reports until 2019 due to the availability of our current set of ESG ratings. To

ensure our dataset is free from survivorship bias, we obtain monthly index compositions from Thomson Reuters Datastream for this period. We consolidate all International Security Identification Numbers (ISINs) and remove any duplicates. Additionally, we exclude financial firms from our dataset (see also [Reverte, 2016](#)). This leads to 5314 firm-year observations without missing control variable data.

We use ESG ratings from different rating agencies, including MSCI KLD, Sustainalytics, Refinitiv, Institutional Shareholder Services (ISS), and Vigeo Eiris. The primary focus is on the MSCI KLD ESG rating, as some textual variables are explicitly derived from the *MSCI KLD ESG Rating Guideline*.

Our dataset includes ESG ratings for 197 firms, resulting in 1618 firm-year observations from 2002 to 2019. Using our word list based on the *MSCI KLD ESG Rating Guideline* (see [section 3.2](#)), we analyze the consistency of ESG disclosure with ESG ratings from MSCI KLD. Out of the 197 firms, only 118 firms have specific ESG ratings from MSCI KLD (resulting in 212 firm-year observations), and these ratings are only available in 2017 and 2018. The period is interesting because German firms are legally required to report on sustainability issues. In addition to our primary ESG ratings, this study incorporates ESG ratings from multiple rating agencies. Our dataset includes 804 firm-year observations from Sustainalytics (2009–2019), 839 firm-year observations from Refinitiv (2002–2019), 925 firm-year observations from Vigeo Eiris (2004–2019), and 1319 firm-year observations from ISS (2006–2019). We show the detailed sample construction in [Appendix A](#).

3.2. Textual data

For the automated textual analysis, we focus on annual reports for several reasons. First, annual reports are recognized as the most reliable disclosure to quantify a firm's contribution to ESG ([Baier et al., 2020](#)). Annual reports contain financially relevant information, are widely recognized by stakeholders including investors (part of our research focus), and have a high degree of credibility. Second, we follow other studies to allow a comparison of the results (e.g., [Bouten et al., 2011](#); [Chan et al., 2014](#); [Giles & Murphy, 2016](#); [Verbeeten et al., 2016](#)). Third, there is a high correlation across sustainability information from different media ([Hooks & van Staden, 2011](#)), i.e., including information sources other than annual reports does not necessarily increase the amount of relevant information. Finally, annual reports include non-financial statements and ESG information, and empirical findings indicate that separate reporting of ESG information has less information value compared to the integrated reporting of ESG information ([Heichl & Hirsch, 2023](#)). We perform automated textual analysis and focus on keywords. We create two word lists. First, we search for existing ESG-related word lists in the literature. We use twelve studies that publish their keywords and extracted keywords from the SDG descriptions, thereby including keywords that are applied in previous studies. The SDGs are precisely presented which allows using the presented keywords for the textual analysis.

Some studies perform manual textual analysis and differentiate between intensities of reporting (e.g., not, briefly, detailed, or extensively reported, see [Wiseman, 1982](#)). We conduct an automated textual analysis using the keywords but not the intensity differentiation.¹¹ For manual analyses that consider whole sentences, we select single keywords or word chains from the longer statements (e.g., 'participation in

environmental organizations' is searched as the combination of 'environment' together with 'organization'). Similar to other studies (e.g., [Baier et al., 2020](#)) we exclude keywords related to quantitative outcomes, e.g., the number of accidents and its comparison with the previous year. We perform a pre-processing of the reports instead of stemming and unify different formulations into a single form for the textual analysis (e.g., changing plural formulations to singular). We include synonyms and eliminate duplicates. Additionally, we exclude very generic keywords. We display a list of excluded words together with our final word list in [Appendix B](#). Although the final academic word list covers 764 keywords, 118 words cannot be found in any report of the sample, e.g., 'migration policy' or 'poaching'. We find hits for 646 words: 241 in the S-pillar, 228 in the E-pillar and 177 in the G-pillar.

Second, we create a word list based on the *MSCI KLD ESG Rating Guideline*. Our unit of analysis is words,¹² so we explicitly search single words and highlight the keywords in each pillar. Similar to the other word list, we only focus on qualitative information. If one single word cannot cover the guideline's requirements, we highlight a word chain. In the next step, we check the results from the highlighting process for duplicates, reduce the number of words in a word chain as much as possible, and we check for overlaps with other non-ESG reporting guidelines. For example, the *MSCI KLD ESG Rating Guideline* values the application of "international standards" but this word chain would be misleading as firms report using the International Financial Reporting Standard (IFRS), reflecting a constant rather than additional content.

For the textual analysis, we employ word vectors, i.e., splitting each document into single words (tokenization). For the word chains, we use different windows in which we expect two coherent words (e.g., setting a base word such as 'global' and adding the term 'warming'). A window of one means the additional word must be immediately before or after the base word (symmetric word window). Setting the window at six means that five other words can be located in between the base and the addition. We test several windows (6, 10, and 14 words) and perform the analysis with a window of 14 words.¹³ Reducing the window keeps the number of words found mainly the same, i.e., the enlargement of the window does not create noise. We manually test several critical formulations and we are aware of double counting (e.g., [Loughran et al., 2009](#)). When considering word chains, we only count the combination of two words instead of counting both single words. For example, we count the word 'water' and the word chain 'water reduction' separately, ensuring that the word 'water' in the chain does not affect the single keyword 'water'. In other words, we are looking for the keyword 'water' only if it is not part of a defined word chain. This differentiation is important because the word lists include different actions or intensities that can relate to the same topic (see [Appendix F](#) and the cited studies following this logic as well).¹⁴

Following [Prasad's \(2008\)](#) guidelines for content analysis, the formulation of categories is important. We follow the original assignment of empirical studies to the E-, S-, and G-pillar as well as into sub-categories. If the study does not differentiate between categories, we assign the categories following the logic used in other studies. For the academic word list, we use the guideline's categorization to assign keywords into pillars and subcategories.

The MSCI word list covers 589 keywords or related word chains, but not every keyword appears in our sample. The final word list only

¹¹ For ESG reporting in Germany, firms are obliged to report about sustainability since 2017, indicating a mandatory reporting setting. However, the content can be voluntarily arranged indicating a voluntary framework as well. Therefore, we assume that firms have incentives to disclose voluntarily information to generate benefits. In contrast, a purely mandatory framework does not allow differentiating between the sustainability engagements in the same way if every firm reports out the same topics. In this case, a content analysis instead of counting words would be necessary.

¹² The unit of analysis reflects the component through which variables are measured (see e.g., [Neuendorf, 2002](#)).

¹³ The results remain constant for both word lists using different word windows. We display the results for the MSCI word list in [Table 3](#).

¹⁴ For this reason, we do not stem our words. We use the pre-processing to reduce the variation of formulations. However, we only want to include the words mentioned in the word lists. When stemming the keywords, there is no guarantee that additional words would not be included. In this case, we could not publish the word lists without concerns.

Table 1
Summary statistics.

Panel A: Firm-year obs. without missing control variable data						
	No.	Median	Mean	SD	Min	Max
<i>Sum_academic</i>	5314	35.48	43.90	29.40	0.31	137.33
<i>Dummy_academic</i>	5314	67	72.93	35.73	0	247
<i>Sum_MSCI</i>	5314	2.96	4.72	5.95	0	65.41
<i>Dummy_MSCI</i>	5314	10	14.07	13.01	0	105
<i>MtB</i>	5314	1.59	1.96	54.93	48.07	1400
<i>Leverage</i>	5314	0.56	0.66	2.23	0.01	83.59
<i>Years_listed</i>	5314	12	14.64	10.33	1	46
Panel B: Firm-year obs. with ESG rating from MSCI KLD						
	No.	Median	Mean	SD	Min	Max
<i>Sum_academic</i>	212	86.04	88.30	32.10	14.02	137.33
<i>Dummies_academic</i>	212	119.5	124.75	38.23	27	243
<i>Sum_MSCI</i>	212	9.65	12.38	10.13	0.97	65.41
<i>Dummies_MSCI</i>	212	30	34.49	18.78	5	97
<i>MtB</i>	212	2.46	3.28	2.65	0.45	14.25
<i>Leverage</i>	212	0.55	0.54	0.17	0.09	0.92
<i>Years_listed</i>	212	18.5	20.37	12.75	1	45
Panel C: Statistics by rating agency						
	No.	Median	Mean	SD	Min	Max
<i>MSCI_Rating</i>	212	14.29	19.33	19.59	7.14	72.14
<i>Sustainalytics</i>	804	57.79	59.48	10.86	37.83	86.25
<i>Refinitiv</i>	839	51.75	51.91	21.95	3.88	94.13
<i>Vigeo Eiris</i>	925	34	34.20	11.77	8	63
<i>ISS</i>	1319	2.17	2.10	0.56	1.06	3.95

Notes: Panel A shows the summary statistics for all firm-year observations without missing control variable data. This includes firms with and without an ESG rating. Panel B shows the summary statistics for all firm-year observations with an ESG rating from MSCI KLD. The period covers the years 2017 and 2018. The disclosure measures base on the MSCI word list or the academic word list. We explain the calculation of the variables in [Appendix B](#). *Sum_academic* is winsorized at 1 % and 99 % levels. Panel C displays the statistics grouped by the mentioned rating agencies. The period covers the years 2002 to 2019. Example: The mean value of *MSCI_Rating* is 19.33 with a standard deviation of 19.59, the median is 14.29 with a range from 7.14 to 72.14.

includes keywords that are disclosed by firms in our sample covers 363 words. The S-pillar contains 190 words, the E-pillar 127, and the G-pillar 46. We present the statistics of both word lists in [Appendix G](#). For both word lists, the distribution of words aligns with the findings from [Dorfleitner et al. \(2015\)](#), showing that the social dimension accounts for most data points and the G-pillar shows a smaller number of data points.

We identify the words that mostly drive the results, following Zipf's law ([Manning & Schütze, 2000](#)), where some words represent a high amount of all hits. For the MSCI word list the top ten mentioned words already reflect 50 % of all hits and the top two words 'safety' and 'fraud' represent 28 % of all hits. In comparison, the academic word list shows a more unified distribution, and the top 41 words reflect 50 % of all hits. Management is likely to repeat relevant ESG terms to highlight specific engagements ([Aldridge & Martin, 2022](#)). Therefore, we compare two textual measures. We count the absolute number of keywords disclosed and thereby include repetitions of keywords (*Sum_MSCI* and *Sum_academic*). This measure is interesting as the repetition of ESG-related keywords can imply a higher intensity of ESG activities. On the other hand, repeating keywords could simply result from longer reports. Previous studies highlight the relationship between firm size and ESG ratings (e.g., [Drempetic et al., 2020](#); [Gamerschlag et al., 2011](#)). Larger firms generally show larger reports due to higher complexity, higher visibility and thereby higher pressure and scrutiny from stakeholder groups ([Brown et al., 2020](#); [Gallo & Christensen, 2011](#)). Therefore we adjust the variables *Sum_MSCI* and *Sum_academic* by dividing them by the log of the length of the annual report, i.e. all words from the report (see also e.g., [Lang & Stice-Lawrence, 2015](#)).

For the second disclosure variable, we use a dummy variable for each

keyword (see e.g., [Clarkson et al., 2013, 2008](#); [Gamerschlag et al., 2011](#); [Plumlee et al., 2015](#)), i.e. if the keyword is reported the dummy is one and zero otherwise. Repetitions have no influence in this variable. As in a checklist, the final measure consists of counting all dummy variables (*Dummies_MSCI* and *Dummies_academic*), and reflects the number of reported keywords in an annual report instead of the sum of the reported keywords. The sum of all single dummy variables for the academic word list can be each value between 0 and 646. It is important to consider both measurements, as we do not know how ESG rating agencies incorporate disclosures into their ratings.

3.3. Regression model

We use panel regression to examine the explanatory power of our ESG disclosure measures and ESG ratings. We use the following base regression model:

$$ESG_Rating_{i,t} = \beta_0 + \beta_1 * ESGD_{i,t} + Controls + \varepsilon_{i,t} \quad (1)$$

ESG_Rating_{i,t} is the rating from different rating agencies in year *t*. The absolute values, including the scale, vary between ESG rating agencies – a higher score implies a better ESG rating. Not all ratings are available for each firm in each year. We explain further details in the [Appendix A and B](#). *ESGD_{i,t}* reflects the placeholder for different disclosure variables from the MSCI and the academic word list, respectively, based on the annual report in year *t*. We include industry-fixed effects and time-fixed effects, as these are commonly addressed external determinants. Selecting control variables is challenging because it is unknown which specific information ESG rating agencies incorporate in the ratings. However, it is known that the focus is on sustainability information and its risks factors ([Billio et al., 2021](#)). Therefore, we decided not to include performance-related measures (e.g., return on assets or sales). We do not include firm size because it will influence the ESG disclosure (see e.g., [Hahn & Kühnen, 2013](#)) as well as the ESG rating (see e.g., [Drempetic et al., 2020](#)) and thereby can add distortions by endogeneity. Furthermore, this will likely introduce firm size-bias to our analysis ([Drempetic et al., 2020](#)).¹⁵ Instead, we add the number of years a firm is listed on the stock market (*Years_listed*), growth opportunities (*MtB*), and leverage. Leverage represents financial risk (e.g., [Gilje, 2016](#); [Hurdle, 1974](#)) that influences sustainable investments and thereby can be considered by ESG rating agencies. Additionally, leverage is related to corporate risk management and managerial risk-taking (e.g., [Kim et al., 2017](#)), which can also be important. The inclusion of growth opportunities is based on the assumption that firms with higher growth prospects could be more likely to engage in sustainable practices to ensure long-term value creation and to meet investor expectations. The number of years a firm has been listed is included as a control variable, as firm age may be related to sustainability practices considered by ESG rating agencies. While ESG practices could evolve over time, firm age itself is an exogenous factor that cannot be influenced by firm decisions. Firm age can shape ESG engagement through mechanisms such as accumulated reputation, organizational learning, and institutional adaptation ([Coad et al., 2018](#)). See [Appendix B](#) for definitions of the variables. We match the ESG disclosure with the ESG rating from the same year *t*.

4. Results and discussion

4.1. Descriptive statistics

In [Table 1](#), we provide the summary statistics of the variables. In Panel A, we display all firm-year observations without missing control

¹⁵ We test the model including firm size and the results show a higher R². This underlines the issue of firm size-bias that strongly influences the variance explained in the ESG rating. We show the results that exclude the firm size effect.

Table 2
Correlation matrices.

Panel A: ESG pillars (MSCI word list)					
	E	S	G	ESG (Dummies_MSCI)	
E	1				
S	0.7551***	1			
G	0.5989***	0.7361***	1		
ESG (Dummies_MSCI)	0.8972***	0.9586***	0.7899***	1	
Panel B: ESG pillars (academic word list)					
	E	S	G	ESG (Sum_academic)	
E	1				
S	0.4900***	1			
G	0.5142***	0.7727***	1		
ESG (Sum_academic)	0.7296***	0.9065***	0.9156***	1	
Panel C: ESG ratings					
	Sustainalytics	Refinitiv	Vigeo Eiris	MSCI KLD	ISS
Sustainalytics	1				
Refinitiv	0.7484***	1			
Vigeo Eiris	0.7612***	0.7456***	1		
MSCI KLD	0.6661***	0.6274***	0.6285***	1	
ISS	0.6573***	0.6169***	0.6735***	0.5103***	1

Note: In this table, we report the Pearson correlation matrices. Panel A and Panel B show correlations for each word list by considering the different pillars and the overall disclosure variable. The values are based on the whole data set (5314 obs.). Panel C shows the correlations between the ESG ratings from all rating agencies (1618 obs.). For example, the correlation coefficient between the ESG rating from ISS and ESG rating from Sustainalytics is 0.6573 with a statistical significance at the 1 % level, meaning that although both measures reflect an ESG rating, the measures are not identical. The correlations range from 0.5103 to 0.7612. The symbols ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels, respectively.

Table 3
Panel regression and comparison of disclosure measures.

Dependent variable = <i>MSCI_Rating</i>								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Dummies_MSCI(6)</i>	0.561*** (0.065)							
<i>Sum_MSCI(6)</i>		0.860*** (0.114)						
<i>Dummies_MSCI(10)</i>			0.524*** (0.060)					
<i>Sum_MSCI(10)</i>				0.839*** (0.110)				
<i>Dummies_MSCI</i>					0.509*** (0.058)			
<i>Sum_MSCI</i>						0.831*** (0.109)		
<i>Sum_academic</i>							0.358*** (0.033)	
<i>Dummies_academic</i>								0.265*** (0.027)
Constant	1.568 (2.193)	9.071*** (1.698)	1.817 (2.133)	9.075*** (1.684)	1.764 (2.137)	9.045*** (1.690)	12.292** (2.714)	13.700*** (3.358)
Obs.	212	212	212	212	212	212	212	212
Adj. R ²	0.230	0.182	0.235	0.185	0.238	0.184	0.344	0.267

Notes: This table displays the regression results for each ESG disclosure variable. We conduct our main analyses with a word window of 14 – the variables are highlighted in bold font; the numbers in the variable names indicate the other tested word windows. The dependent variable is the ESG rating from MSCI KLD. The sample covers the years 2017 and 2018. For a regression analysis with the ESG rating from MSCI KLD as dependent variable, for example, the regression coefficients for *Dummies_MSCI*, and *Sum_academic* are 0.509 and 0.358, respectively. We use robust standard errors. The symbols ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels, respectively; t-statistics are displayed in parentheses. All variables are explained in [Appendix B](#).

variable data. This sample includes firms with and without an ESG rating between 2002 and 2019. In Panel B, we focus on the firms that received an MSCI KLD ESG rating (covers the years 2017 and 2018). Firms with the ESG rating show higher mean and median of the disclosure variables compared to the overall sample in Panel A. In Panel C, we provide the statistics of the ESG ratings.

We graphically report the development in the disclosure scores over time in [Appendix C](#). The figure shows the continuous rise in ESG disclosure. We report descriptive statistics and divide the sample into percentiles based on the disclosure variables (*Dummies_MSCI*

and *Sum_academic*) and the ESG rating from MSCI KLD. The results provided in [Appendix D and E](#) reveal that firms with higher ESG disclosure have a significantly higher ESG rating, higher leverage, and longer stock market listing. The characteristics are consistent across both disclosure variables and similar among the ESG ratings.

The descriptive statistics are an initial assessment of the ESG variables' quality and illustrate similarities with existing studies. [Bonacorsi et al. \(2024\)](#) find for their international sample covering ESG ratings from MSCI KLD that a substantial group of companies shows only 150 ESG entries out of all possible 537 potential variables. [Baier et al. \(2020\)](#)

Table 4
Panel regression and comparison of disclosure measures including fixed-effects and controls.

Dependent variable = <i>MSCI_Rating</i>								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Dummies_MSCI(6)</i>	0.430*** (0.079)							
<i>Sum_MSCI(6)</i>		0.646*** (0.116)						
<i>Dummies_MSCI(10)</i>			0.409*** (0.073)					
<i>Sum_MSCI(10)</i>				0.632*** (0.114)				
<i>Dummies_MSCI</i>					0.399*** (0.071)			
<i>Sum_MSCI</i>						0.625*** (0.113)		
<i>Sum_academic</i>							0.309*** (0.040)	
<i>Dummies_academic</i>								0.209*** (0.033)
Constant	0.571 (4.613)	3.744 (4.364)	0.291 (4.536)	3.790 (4.357)	0.379 (4.521)	3.748 (4.361)	9.453** (4.150)	11.883** (5.212)
Industry & time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	212	212	212	212	212	212	212	212
Adj. R ²	0.265	0.255	0.268	0.256	0.271	0.255	0.387	0.288

Notes: This table displays the regression results for each ESG disclosure variable. We conduct our main analyses with a word window of 14 – the variables are highlighted in bold font; the numbers in the variable names indicate the other tested word windows. The dependent variable is the ESG rating from MSCI KLD. The sample covers the years 2017 and 2018. For a regression analysis with the ESG rating from MSCI KLD as dependent variable, for example, the regression coefficients for *Dummies_MSCI*, and *Sum_academic* are 0.399 and 0.309, respectively. We use robust standard errors. The symbols ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels, respectively; t-statistics are displayed in parentheses.

Table 5
Panel regression for ESG pillars.

Dependent variable = <i>MSCI_Rating</i>								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>E_MSCI</i>	1.178*** (0.165)							
<i>S_MSCI</i>		0.853*** (0.105)						
<i>G_MSCI</i>			2.435*** (0.462)					
<i>Dummies_MSCI</i>				0.509*** (0.058)				
<i>Sum_E_academic</i>					0.629*** (0.111)			
<i>Sum_S_academic</i>						0.613*** (0.052)		
<i>Sum_G_academic</i>							0.583*** (0.056)	
<i>Sum_academic</i>								0.358*** (0.033)
Constant	8.868*** (1.717)	2.462 (2.240)	5.128* (2.737)	1.764 (2.137)	11.744*** (1.610)	3.839* (2.085)	6.448*** (2.287)	12.292*** (2.714)
Obs.	212	212	212	212	212	212	212	212
Adj. R ²	0.216	0.208	0.099	0.238	0.189	0.310	0.311	0.344
p-Value Ramsey Test	0.2659	0.7062	0.2172	0.7277	0.000	0.4285	0.0961	0.1691

Notes: This table shows the impact of each ESG pillar individually as well as the overall ESG disclosure variables. We use a word window of 14. The dependent variable is the ESG rating from MSCI KLD. All variables are explained in detail in [Appendix B](#). The sample covers the years 2017 and 2018. We add Ramsey Reset Test ([Ramsey, 1969](#)) to control for omitted variables. For a regression analysis with the ESG rating from MSCI KLD as dependent variable, for example, the regression coefficients for *Dummies_MSCI*, and *Sum_academic* are 0.509 and 0.358, respectively. We use robust standard errors. The symbols ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels, respectively; t-statistics are displayed in parentheses.

report an overall share of ESG-related information at about 3.7 %, which is higher than in our sample.¹⁶ This can be attributed to their sample consisting of the 10-K reports of the 25 largest companies in the S&P100 index and their textual analysis procedure.

We present a correlation matrix for the different ESG pillars and ESG

ratings in [Table 2](#), with separate panels for each word list. The results reveal differences between the two word lists. Consistent with the findings of [Sahin et al. \(2022\)](#), the E- and S-pillars in the MSCI word list show stronger correlations with both the overall ESG disclosure variable (i.e., *Dummies_MSCI*) and with each other than with the G-pillar. In contrast, for the academic word list, the G-pillar exhibits the highest correlation with the overall ESG disclosure variable (i.e., *Sum_academic*). The correlations between the E- and G-pillars, as well as

¹⁶ We have an overall share of ESG-related information of 1.59 %.

between the S- and G-pillars, are similar across both word lists. Panel C reports the correlations between ESG ratings from different rating agencies, illustrating the well-documented issue of rating dispersion. Although all scores aim to capture a firm's ESG activity, the pairwise correlations range from 0.51 to 0.76, indicating notable variation among providers.

4.2. Regression analysis

As explained above, we create two word lists and, for each, two overall disclosure measures. We differentiate between the absolute word count, which includes repetitions (*Sum_MSCI* and *Sum_academic*) and the presence of keywords, captured as dummy variables for each keyword. The sum of all single dummy variables is considered as the second overall disclosure measure (*Dummies_MSCI* and *Dummies_academic*).

In the first step, we present the selection of disclosure variables and analyze them individually using different word window sizes.¹⁷ Table 3 reports the results. The word window used for each variable is indicated in parentheses; if not specified, a window size of 14 is applied. The dependent variable across all models is the ESG rating from MSCI KLD. To control for outliers, we winsorize the *Sum_academic* variable at the 1% and 99% level.

Models (1) to (6) focus on the MSCI word list. Specifically, models (1), (3), and (5) use the sum of dummy variables, while models (2), (4), and (6) use the absolute word count (sum variables). The adjusted R^2 values are relatively consistent across different word windows, with the window of 14 yielding the highest explanatory power. However, given the minimal variation across windows, the choice of 14 does not appear to introduce noise and is used in subsequent analyses.

Comparing the sum of dummy models (1, 3, 5) to the sum models (2, 4, 6), we find that the dummy variables explain a higher proportion of the variation in MSCI KLD ratings. This suggests that the presence of ESG-related terms – rather than their frequency – is more informative when using the MSCI word list. One interpretation is that extensive repetition of keywords does not enhance a firm's rating, potentially indicating that ESG raters are not misled by overly detailed disclosures (i.e., limited greenwashing effect). Alternatively, this finding may reflect MSCI's use of binary indicators in its rating methodology (Dorfleitner et al., 2015). Based on these results, we continue using the *Dummies_MSCI* variable for further analysis.

Models (7) and (8) use the academic word list. Both models apply a word window of 14, as variations across window sizes again showed minimal impact. In contrast to the MSCI word list, the *Sum_academic* variable (model 7) shows a higher adj. R^2 than the dummy-based variable (model 8). Furthermore, both academic disclosure variables exhibit stronger explanatory power than any of the disclosure variables from the MSCI word list (models 1 to 6). This suggests that the academic word list aligns more closely with the components of the MSCI KLD rating.

We test our results by including controls and industry- and time-fixed-effects. The results in Table 4 show that the adj. R^2 increases in each model. However, the academic disclosure variables still have higher explanatory power compared to the disclosure variables based on the MSCI word list. The highest variance is explained by *Sum_academic*. Based on these results, we continue the further analyses with the variables that show the highest explanatory power, i.e., the *Sum_academic* variable (adj. R^2 : 0.387) and the *Dummies_MSCI* variable (adj. R^2 : 0.271).

We divide our sample into the three pillars E, S, and G and further subcategories. Previous studies find different results for different pillars (e.g., Berg, Heeb, & Kölbl, 2022; Bird et al., 2007; Galema & Gerritsen, 2025; Khan et al., 2016; Kimbrough et al., 2024; Schiemann & Tietmeyer, 2022; Verbeeten et al., 2016). Table 5 models 1 to 4 show

positive and highly significant coefficients for each pillar using the MSCI variables. A low explanatory power is caused by the G-pillar, and the highest by the E-pillar. Using the academic word list in models 5 to 8 shows a higher explanatory power for the S- and G-pillar. The results remain robust when including control variables. Contrary to our assumption, we do not find a higher consistency between the MSCI word list and the MSCI KLD ESG rating compared to the academic word list. This may suggest that ESG rating agencies, even when publishing guideline-based frameworks, incorporate additional qualitative judgments or external data sources not fully reflected in narrative disclosures.

Interestingly, the overall disclosure variables (models 4 and 8) have only slightly higher explanatory power than the single E-pillar of the MSCI word list and the S- or G-pillar of the academic word list (models 1 and 6). These findings reinforce the notion that the explanatory power of ESG disclosure varies across the three pillars. Interestingly, when using the academic word list, the S- and G-pillars display stronger explanatory power than with the MSCI word list. This suggests that the academic word list may capture a broader or more nuanced set of words relevant to these dimensions.

We test for omitted variables and use the Ramsey Reset test (Ramsey, 1969). The test indicates that there is no specification error in the data (p-value for F-stat is 0.1691 for *Sum_academic* variable; see Table 5, model 8). Including control variables, the p-value is significant at the 10% level. The effect is driven by industry-fixed effects and the results are robust when excluding the industry fixed-effects (p-value for F-stat is 0.2246 for *Sum_academic* variable; see Table 6, model 9) which means we include industry-fixed effects for further models without a specification error.

Besides the MSCI universe, we use additional ESG ratings and therefore repeat our analysis. We include the *Dummies_academic* variable to be able to detect differences between ESG rating providers. The *Dummies_academic* variable still relates to the lowest model fit while *Sum_academic* explains a higher proportion of the variance in the underlying rating compared to the *Dummies_MSCI* for several ratings. Table 7 shows the results. For most of the ESG ratings, the disclosure scores including repetitions of keywords show a higher model fit. However, this is not the case for the ESG rating from Vigeo Eiris (Panel B, model 4 and 6). *Dummies_MSCI* explains a higher proportion of the variance in the ESG rating compared to the academic variables, highlighting the need for caution when conducting textual analysis. We explain the differences in the results by the effect of ESG rating dispersion (Berg, Heeb, & Kölbl, 2022). For the *Sum_academic* variable, the highest model fit is for Refinitiv ratings (adj. R^2 : 0.447), while the lowest model fit relates to the ISS ratings (adj. R^2 : 0.140). This finding suggests that ISS likely relies on other information besides the annual report or employs more sophisticated methods in its analysis.

To understand the influence of different components, we consider the subcategories in Table 8 and Table 9. For this analysis, we divide the sum of all words for subcategories by the number of keywords in the subcategory to exclude the effects from the keyword distribution. While each ESG pillar is statistically significant, this is not the case for the subcategories. For the academic word list (Table 9), we find that the subcategories 'environmental opportunities' and 'natural capital' are not significant. Based on the MSCI word list (Table 8) the subcategories from the G-pillar show insignificant coefficients or a reduced level of significance. Subcategories with the highest model fit are 'product liabilities' and 'climate change' for the MSCI word list and 'human capital', 'corporate governance', and 'climate change' for the academic word list. Besides the insignificant coefficients in Panel A, based on the academic word list, the model fit for 'social opportunities' and 'stakeholder opposition' is lower compared to the results based on the MSCI word list. The controversial subcategory reveals interesting results. While there is a negative and significant coefficient for the academic word list when including controls and fixed effects, the MSCI word list shows a positive coefficient (significant at 5% level). Given that we find the academic

¹⁷ The settings for word windows are explained in Section 3.2.

Table 6
Panel regression for ESG pillars including fixed-effects and controls.

Dependent variable = <i>MSCI_Rating</i>	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<i>E_MSCI</i>	0.955*** (0.190)								
<i>S_MSCI</i>		0.629*** (0.130)							
<i>G_MSCI</i>			1.334*** (0.513)						
<i>Dummies_MSCI</i>				0.399*** (0.071)					
<i>Sum_E_academic</i>					0.455*** (0.124)				
<i>Sum_S_academic</i>						0.518*** (0.060)			
<i>Sum_G_academic</i>							0.490*** (0.065)		
<i>Sum_academic</i>								0.309*** (0.040)	0.321*** (0.038)
Constant	3.961 (4.343)	0.339 (4.597)	1.730 (4.473)	0.379 (4.521)	6.802 (4.132)	3.677 (4.094)	5.846 (3.994)	9.453** (4.150)	11.529 (4.005)
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	212	212	212	212	212	212	212	212	212
Adj. R ²	0.263	0.256	0.192	0.271	0.242	0.338	0.336	0.387	0.349
p-Value Ramsey Test	0.0584	0.7779	0.5385	0.4523	0.1492	0.5534	0.0056	0.0768	0.2246

Notes: This table shows the impact of each ESG pillar individually as well as the overall ESG disclosure variables. We use a word window of 14. The dependent variable is the ESG rating from MSCI KLD. All variables are explained in detail in [Appendix B](#). The sample covers the years 2017 and 2018. We add Ramsey Reset Test ([Ramsey, 1969](#)) to control for omitted variables. For a regression analysis with the ESG rating from MSCI KLD as dependent variable, for example, the regression coefficients for *Dummies_MSCI*, and *Sum_academic* are 0.399; and 0.309, respectively. We use robust standard errors. The symbols ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels, respectively; t-statistics are displayed in parentheses.

word list to show overall a higher explanatory power for several ratings, the results help to identify which subcategory of the MSCI word list performs better.

4.3. Robustness checks and discussion

In this section, we perform several robustness checks. We include quarterly reports and stand-alone sustainability reports to enhance the number of firms' communication channels and conduct a principal component analysis.

First, we add quarterly reports to our sample (see [Tables 10 and 11](#)). We do not simply add the observations from quarterly reports to the regression model because the control variable data is on a yearly base, which would lead to distortions. Instead, we create new variables by adding the number of reported items (i.e., dummies for reported words or sum of reported words, respectively) from the quarterly reports to the annual report observations. Consequently, the number of firm-year observations is not different from the main analysis. The results vary between the ESG rating providers, with slight changes due to the inclusion of quarterly report information. For the ESG ratings from ISS, Vigeo Eiris, and Sustainalytics we find an increase in the adjusted R² of the *Dummies_MSCI* compared to using only annual reports (adj. R² ISS: 0.137 with quarterly reports vs. 0.130 without; adj. R² Vigeo Eiris: 0.406 with vs. 0.390 without quarterly reports; adj. R² Sustainalytics: 0.353 with vs. 0.331 without quarterly reports). Interestingly, we find a slight reduction in the explanatory power for Refinitiv ratings (adj. R² of 0.419 with vs. 0.430 without quarterly reports). For the MSCI KLD rating, the model fit decreases when considering the MSCI word list but increases slightly when using the academic word list (by 0.011).¹⁸ The differences are smaller for the academic word list. The highest increase in the model fit can be seen for the Sustainalytics rating (increase of 0.022). Overall, the

¹⁸ The adj. R² can be extracted from [Table 10](#), model (4) each panel and [Table 6](#), models (4) and (8).

differences are minor, supporting our decision to focus on annual reports while recommending the inclusion of additional information channels. Quarterly reports may contain environmentally and socially important information, e.g., after special events or shocks.

Besides quarterly reports, we focus on stand-alone ESG reports instead of annual reports following other empirical studies. [Balluchi et al. \(2021\)](#) find that stand-alone sustainability reports are positively related to their credibility measure that covers for example the dimensions of understandability and reliability. Due to limited availability of stand-alone sustainability reports, our sample includes only 103 firms and 517 sustainability reports between 2002 and 2019 with the highest number of reports (94) in 2019. We repeat the analysis conducted for the main study.

First, we focus on the ESG rating from MSCI KLD and divide the disclosure scores by the three pillars (see [Table 12](#)). The results show the highest model fit for the E-pillar using both word lists. Furthermore, the *Sum_academic* variable still shows a higher model fit compared to *Dummies_MSCI*. The G-pillar is only slightly significant using the MSCI word list. This could reflect a lower emphasis on governance-related language in sustainability reports, which often focus more heavily on environmental and social topics. However, the G-pillar remains significant when repetitions are included (Panel B model 3).

In [Table 13](#), we report the results for the other ESG rating agencies and compare them to the results based on annual reports (see [Tables 6 and 7](#)). The coefficients for *Dummies_MSCI* are insignificant for three out of four rating agencies and show only a reduced level of significance for the Sustainalytics rating.

When examining the model fit of the *Sum_academic* variable, the results are mixed. For Vigeo Eiris, the model fit is nearly identical whether the stand-alone sustainability report or the annual report is used ([Table 13](#), model 4 vs. [Table 7](#), Panel B, model 6). Interestingly, the model fit for ISS improves when using the sustainability report compared to the annual report ([Table 13](#), model 2 vs. [Table 7](#), Panel B, model 3). For the remaining ratings, including MSCI KLD, the explanatory power is reduced when switching from the annual report to the

Table 7
Panel regression and comparison of disclosure measures – other ESG ratings.

Panel A: Without control variable data												
Dependent variable =	ISS			Vigeo Eiris			Sustainalytics			Refinitiv		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
<i>Dummies_MSCI</i>	0.007*** (0.002)			0.348*** (0.032)			0.321*** (0.034)			0.700*** (0.064)		
<i>Dummies_academic</i>		0.003*** (0.001)			0.151*** (0.019)			0.152*** (0.017)			0.316*** (0.033)	
<i>Sum_academic</i>			0.004*** (0.001)			0.181*** (0.024)			0.174*** (0.024)			0.393*** (0.042)
Constant	1.896*** (0.084)	1.783*** (0.116)	1.807*** (0.097)	22.964*** (1.512)	15.910*** (2.540)	19.106*** (2.333)	48.587*** (1.453)	40.636*** (2.388)	44.790*** (2.242)	28.191*** (2.882)	12.430*** (4.667)	18.179*** (4.176)
Obs.	1319	1319	1319	925	925	925	804	804	804	839	839	839
Adj. R ²	0.039	0.030	0.048	0.264	0.240	0.252	0.239	0.240	0.233	0.323	0.302	0.352
Panel B: With control variable data and fixed-effects												
Dependent variable =	ISS			Vigeo Eiris			Sustainalytics			Refinitiv		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
<i>Dummies_MSCI</i>	0.010*** (0.002)			0.321*** (0.042)			0.283*** (0.044)			0.572*** (0.077)		
<i>Dummies_academic</i>		0.004*** (0.001)			0.130*** (0.022)			0.129*** (0.023)			0.246*** (0.039)	
<i>Sum_academic</i>			0.005*** (0.001)			0.160*** (0.026)			0.154*** (0.030)			0.333*** (0.050)
Constant	1.853*** (0.138)	1.727*** (0.156)	1.775*** (0.144)	19.651*** (2.870)	14.528*** (3.585)	17.936*** (3.215)	48.548*** (2.526)	42.546*** (3.204)	46.397*** (2.941)	30.251*** (5.385)	19.715*** (6.558)	23.818*** (6.022)
Industry & time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	1319	1319	1319	925	925	925	804	804	804	839	839	839
Adj. R ²	0.130	0.112	0.140	0.390	0.356	0.371	0.331	0.326	0.338	0.430	0.406	0.447

Notes: This table shows separate regressions: The dependent variables are the ESG ratings from ISS; Vigeo Eiris, Sustainalytics, and Refinitiv, respectively. Panel B includes control variable data as well as fixed effects. For a regression analysis with the ESG rating from Refinitiv as dependent variable (Panel B, model 10 to 12), for example, the regression coefficients for *Dummies_MSCI*, and *Sum_academic* are 0.572 and 0.333, respectively. We use robust standard errors clustered by firm. The symbols ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels, respectively; t-statistics are displayed in parentheses.

Table 8
Panel regression with different subcategories for each ESG pillar (MSCI word list).

Dependent variable = <i>MSCI_Rating</i>								
Panel A: E-Pillar								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Dummies_climate</i>	2.593*** (0.376)	1.978*** (0.416)						
<i>Dummies_env_opp</i>			3.589*** (0.601)	2.695*** (0.612)				
<i>Dummies_naturalcapital</i>					3.374*** (0.606)	2.303*** (0.738)		
<i>Dummies_pollution</i>							7.159*** (1.357)	5.286*** (1.528)
Constant	9.800*** (1.595)	4.316 (4.511)	10.845*** (1.716)	3.949 (4.245)	13.868*** (1.512)	5.468 (4.373)	15.579*** (1.432)	6.474 (4.383)
Industry and time FE		Yes		Yes		Yes		Yes
Controls		Yes		Yes		Yes		Yes
Obs.	212	212	212	212	212	212	212	212
Adj. R ²	0.183	0.239	0.153	0.235	0.126	0.210	0.136	0.232
Panel B: S-Pillar								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Dummies_human</i>	1.551*** (0.221)	1.114*** (0.267)						
<i>Dummies_social_opp</i>			7.064*** (1.154)	5.377*** (1.331)				
<i>Dummies_product</i>					2.086*** (0.283)	1.548*** (0.297)		
<i>Dummies_stakeholder</i>							6.356*** (1.115)	3.855*** (1.402)
Constant	4.770** (2.318)	0.226 (4.666)	12.763*** (1.564)	5.002 (4.292)	2.353 (2.607)	1.107 (4.494)	12.492*** (1.620)	6.176 (4.267)
Industry and time FE		Yes		Yes		Yes		Yes
Controls		Yes		Yes		Yes		Yes
Obs.	212	212	212	212	212	212	212	212
Adj. R ²	0.161	0.239	0.143	0.238	0.203	0.261	0.140	0.205
Panel C: G-Pillar								
	(1)	(2)	(3)	(4)	(5)	(6)		
<i>Dummies_behaviour</i>	1.603** (0.620)	0.084 (0.663)						
<i>Dummies_governance</i>			4.879*** (1.613)	2.300 (1.856)				
<i>Dummies_controversial</i>					11.183*** (3.565)	8.260** (3.640)		
Constant	14.109*** (2.445)	6.488 (4.460)	10.535*** (2.832)	3.298 (4.567)	17.692*** (1.377)	6.650 (4.314)		
Industry and time FE		Yes		Yes		Yes		
Controls		Yes		Yes		Yes		
Obs.	212	212	212	212	212	212		
Adj. R ²	0.024	0.166	0.043	0.175	0.053	0.193		

Notes: This table shows the subcategories of the ESG-pillars. The dependent variable is the MSCI KLD ESG rating. The textual variables are based on the MSCI word list; we use a word window of 14. All variables are explained in detail in Appendix B. The sample covers the years 2017 and 2018. For a regression analysis based on the MSCI word list with the ESG rating from MSCI KLD as dependent variable including control variables and fixed effects, for example, the regression coefficients of the subcategories *Dummies_pollution*, and *Dummies_human* are 5.286 and 1.114, respectively. We use robust standard errors. The symbols ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels, respectively; t-statistics are displayed in parentheses.

sustainability report. In the case of Refinitiv, the *Sum_academic* variable is only marginally significant when based on the sustainability report.

These findings suggest that ESG rating agencies may assess stand-alone sustainability reports more critically than annual reports. However, ISS appears to place greater weight on the information contained in sustainability reports. The insignificance of the *Dummies.MSCI* variable further indicates that firms may need to mention ESG-related keywords multiple times to signal their ESG engagement effectively. Beyond the results, it is important to note that annual reports remain a key source of ESG information, as they are consistently available across firms, whereas stand-alone sustainability reports are not.

Finally, we perform an unsupervised machine learning approach using a principal component analysis (PCA) to select words by their magnitude and reducing the amount of keywords while maintaining a high information value. This approach can improve the model fit by

including relationships between keywords. PCA is an orthogonal linear transformation through scalar projection. The greatest variance of the data – capturing both between-firm differences and within-firm variation over time – is reflected in the first coordinate (principal component) while the second coordinate explains most of the variance when the first is removed (see also Bonacorsi et al., 2024). An important condition for PCA is standardizing the variables (Abdi & Williams, 2010). PCA can be performed using standardized data without distributional assumptions (e.g., Jolliffe & Cadima, 2016), but normal-like data can be helpful for the interpretation of the underlying data. We test both approaches and find robust results. We report results using Box-Cox transformation for a more normal-like distribution, and we use standardized variables. Following Kaiser’s rule, we include principal components with an eigenvalue greater than one (e.g., Jackson, 1993). Instead of focusing on the sum of keywords, we now consider the principal components in the

Table 9
Panel regression with different subcategories for each ESG pillar (academic word list).

Dependent variable = <i>MSCI_Rating</i>								
Panel A: E-Pillar								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Sum_climate</i>	7.084*** (1.137)	5.238*** (1.074)						
<i>Sum_env_opp</i>			1.298 (1.194)	0.957 (1.032)				
<i>Sum_naturalcapital</i>					4.928 (3.286)	0.412 (2.034)		
<i>Sum_pollution</i>							18.182*** (3.597)	13.394*** (3.219)
Constant	11.434*** (1.573)	6.863* (4.116)	18.855*** (1.367)	6.880 (4.280)	18.277*** (1.447)	6.654 (4.381)	13.796*** (1.422)	5.129 (4.152)
Industry and time FE		Yes		Yes		Yes		Yes
Controls		Yes		Yes		Yes		Yes
Obs.	212	212	212	212	212	212	212	212
Adj. R ²	0.212	0.262	0.008	0.171	0.025	0.167	0.157	0.237
Panel B: S-Pillar								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Sum_human</i>	10.359*** (1.000)	8.829*** (1.043)						
<i>Sum_social_opp</i>			86.421*** (25.002)	52.300** (24.301)				
<i>Sum_product</i>					17.177*** (2.133)	13.130*** (1.748)		
<i>Sum_stakeholder</i>							1.055*** (0.220)	0.627*** (0.219)
Constant	1.968 (1.795)	0.134 (4.061)	16.025*** (1.647)	6.376 (4.324)	4.994*** (1.874)	0.490 (4.184)	7.410*** (2.586)	2.778 (4.509)
Industry and time FE		Yes		Yes		Yes		Yes
Controls		Yes		Yes		Yes		Yes
Obs.	212	212	212	212	212	212	212	212
Adj. R ²	0.321	0.351	0.042	0.180	0.183	0.268	0.075	0.188
Panel C: G-Pillar								
	(1)	(2)	(3)	(4)	(5)	(6)		
<i>Sum_behavior</i>	8.789*** (1.234)	6.333*** (1.416)						
<i>Sum_governance</i>			7.130*** (0.723)	6.034*** (0.816)				
<i>Sum_controversial</i>					11.865 (7.816)	16.398*** (4.620)		
Constant	5.287** (2.042)	4.131 (4.203)	5.709** (2.348)	6.309 (3.993)	19.411*** (1.357)	6.706 (4.369)		
Industry and time FE		Yes		Yes		Yes		
Controls		Yes		Yes		Yes		
Obs.	212	212	212	212	212	212		
Adj. R ²	0.220	0.249	0.309	0.347	0.002	0.170		

Notes: This table shows the subcategories of the three ESG-pillars. The dependent variable is the MSCI KLD ESG rating. The textual variables are based on the academic word list; we use a word window of 14. All variables are explained in detail in [Appendix B](#). The sample covers the years 2017 and 2018. For a regression analysis based on the academic word list with the ESG rating from MSCI KLD as dependent variable including control variables and fixed effects, for example, the regression coefficients of the subcategories *Sum_pollution*, and *Sum_human* are 13.394 and 8.829, respectively. We use robust standard errors. The symbols ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels, respectively; t-statistics are displayed in parentheses.

regression model. Thereby, we reduce the number of variables to the number of principal components while preserving as much information as possible. The results do not allow identifying influential keywords, instead the coefficients are constructed as linear combinations and no longer have real meaning ([Bro & Smilde, 2014](#)). Using PCA for the academic word list leads to 157 principal components, explaining 70.6 % of the variance in the MSCI KLD ESG rating. Incorporating principal components into the regression model increases the explanatory power for several ESG ratings. We provide the results in [Table 14](#); due to the amount of principal components, we show only the first and the last component.

The increase in the explanatory power indicates that considering correlations between keywords enhances consistency between the disclosure and different ESG ratings. This aligns with our assumption that weighting different subcategories, keywords, or pillars improves the model fit.

The nature of ESG disclosure is at least in parts qualitative, complicating the comparison between firms ([Amel-Zadeh & Serafeim, 2018](#)). At the same time, investors respond favorably to qualitative disclosure ([Martin & Moser, 2016](#)). Our study focuses on qualitative information, but we have some limitations, as we do not include the evaluation of quantitative measures (e.g., inclusion of ‘work accident’ but not number

Table 10
Panel regression including quarterly reports.

Dependent variable = <i>MSCI_Rating</i>					
Panel A: Word list MSCI					
	(1)	(2)	(3)	(4)	
<i>E_MSCI</i>	0.670*** (0.164)				
<i>S_MSCI</i>		0.550*** (0.105)			
<i>G_MSCI</i>			1.139*** (0.384)		
<i>Dummies_MSCI</i>				0.323*** (0.059)	
Constant	4.269 (4.403)	0.129 (4.507)	1.253 (4.601)	0.001 (4.497)	
Quarterly reports	Yes	Yes	Yes	Yes	
Industry and time FE	Yes	Yes	Yes	Yes	
Controls	Yes	Yes	Yes	Yes	
Obs.	212	212	212	212	
Adj. R ²	0.236	0.262	0.201	0.267	
Panel B: Academic word list					
	(1)	(2)	(3)	(4)	(5)
<i>Sum_E_academic</i>	0.332*** (0.100)				
<i>Sum_S_academic</i>		0.476*** (0.054)			
<i>Sum_G_academic</i>			0.428*** (0.061)		
<i>Sum_academic</i>				0.235*** (0.027)	
<i>Dummies_academic</i>					0.107*** (0.024)
Constant	7.332* (4.135)	3.673 (4.081)	5.003 (4.040)	4.579 (3.859)	5.673 (5.302)
Quarterly reports	Yes	Yes	Yes	Yes	Yes
Industry and time FE	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
Obs.	212	212	212	212	212
Adj. R ²	0.228	0.343	0.321	0.361	0.242

Notes: This table shows the impact of each ESG pillar individually as well as the overall ESG disclosure variables including quarterly reports. We adjust the firm-year observations with respect to the quarterly reports and therefore do not have a higher number of observations. We use a word window of 14. The dependent variable is the ESG rating from MSCI KLD. All variables are explained in detail in Appendix B. The sample covers the years 2017 and 2018. For a regression analysis based on the MSCI word list with the ESG rating from MSCI KLD as dependent variable, including quarterly reports, controls, and fixed-effects, for example, the regression coefficient of *Dummies_MSCI*, is 0.323. We use robust standard errors. The symbols ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels, respectively; t-statistics are displayed in parentheses.

of accidents), although it is considered by ESG rating agencies. A manual classification of words and sentences would allow extracting reporting intensity and context, but it would restrict the analysis to small sample sizes and require major effort. Our word list cannot include all components considered for ESG rating evaluation, as they are not publicly available. The rating methodology, including weights and the data sources, is viewed as intellectual property of rating agencies. We cannot fully replicate the weighting and specific criteria applied in the categories, including evaluation of strengths and concerns. The MSCI KLD methodology incorporates both strengths and weaknesses, but we cannot differentiate between them automatically due to the dependency on the textual content. Although there are different word lists (Loughran

Table 11
Panel regression including quarterly reports – other ESG ratings.

Dependent variable = <i>Refinitiv</i>				
	(1)	(2)	(3)	(4)
<i>Dummies_MSCI</i>	0.572*** (0.077)	0.439*** (0.066)		
<i>Sum_academic</i>			0.333*** (0.050)	0.222*** (0.037)
Constant	30.251*** (5.385)	31.570*** (5.370)	23.818*** (6.022)	29.324*** (5.376)
Controls and FE	Yes	Yes	Yes	Yes
Quarterly reports		Yes		Yes
Obs.	839	839	839	839
Adj. R ²	0.430	0.419	0.447	0.423
Dependent variable = <i>Sustainalytics</i>				
	(1)	(2)	(3)	(4)
<i>Dummies_MSCI</i>	0.283*** (0.044)	0.234*** (0.036)		
<i>Sum_academic</i>			0.154*** (0.030)	0.110*** (0.019)
Constant	48.548*** (2.526)	48.692*** (2.417)	46.397*** (2.941)	48.185*** (2.568)
Controls and FE	Yes	Yes	Yes	Yes
Quarterly reports		Yes		Yes
Obs.	804	804	804	804
Adj. R ²	0.331	0.353	0.338	0.339
Dependent variable = <i>Vigeo Eiris</i>				
	(1)	(2)	(3)	(4)
<i>Dummies_MSCI</i>	0.321*** (0.042)	0.261*** (0.033)		
<i>Sum_academic</i>			0.160*** (0.026)	0.118*** (0.018)
Constant	19.651*** (2.870)	19.783*** (2.744)	17.936*** (3.215)	19.505*** (2.900)
Controls and FE	Yes	Yes	Yes	Yes
Quarterly reports		Yes		Yes
Obs.	925	925	925	925
Adj. R ²	0.390	0.406	0.371	0.376
Dependent variable = <i>ISS</i>				
	(1)	(2)	(3)	(4)
<i>Dummies_MSCI</i>	0.010*** (0.002)	0.008*** (0.002)		
<i>Sum_academic</i>			0.005*** (0.001)	0.004*** (0.001)
Constant	1.853*** (0.138)	1.862*** (0.135)	1.775*** (0.144)	1.839*** (0.138)
Controls and FE	Yes	Yes	Yes	Yes
Quarterly reports		Yes		Yes
Obs.	1319	1319	1319	1319
Adj. R ²	0.130	0.137	0.140	0.140

Notes: This table shows the ESG disclosure variables including the quarterly reports in model (2) and model (4). We adjust the firm-year observations with respect to the quarterly reports and therefore do not have a higher number of observations. The dependent variables are the ESG ratings from the mentioned ESG rating agencies. For a regression analysis with the ESG rating from ISS, including quarterly reports, for example, the regression coefficients of *Dummies_MSCI* and *Sum_academic*, are 0.008 and 0.004, respectively. We use robust standard errors clustered by firm. The symbols ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels, respectively; t-statistics are displayed in parentheses.

& McDonald, 2011), applying them is difficult, as identifiers for strengths and weaknesses differ in the ESG-context compared to, for example, earnings call documents. As Loughran & McDonald, 2016

Table 12
Panel regression based on sustainability reports.

Dependent variable = <i>MSCI_Rating</i>					
Panel A: Word list MSCI					
	(1)	(2)	(3)	(4)	
<i>E_MSCI</i>	0.825*** (0.165)				
<i>S_MSCI</i>		0.404*** (0.130)			
<i>G_MSCI</i>			1.089* (0.604)		
<i>Dummies_MSCI</i>				0.247*** (0.067)	
Constant	3.800 (6.109)	5.834 (6.433)	8.354 (6.700)	4.804 (6.385)	
Industry and time FE	Yes	Yes	Yes	Yes	
Controls	Yes	Yes	Yes	Yes	
Obs.	129	129	129	129	
Adj. R ²	0.314	0.248	0.206	0.268	
Panel B: Academic word list					
	(1)	(2)	(3)	(4)	(5)
<i>Sum_E_academic</i>	0.694*** (0.097)				
<i>Sum_S_academic</i>		0.306*** (0.085)			
<i>Sum_G_academic</i>			0.423*** (0.095)		
<i>Sum_academic</i>				0.184*** (0.033)	
<i>Dummies_academic</i>					0.085*** (0.033)
Constant	2.840 (5.321)	3.920 (6.350)	0.518 (6.122)	0.539 (6.022)	4.679 (6.808)
Industry and time FE	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
Obs.	129	129	129	129	129
Adj. R ²	0.385	0.287	0.304	0.352	0.231

Notes: This table shows the impact of each ESG pillar individually as well as the overall ESG disclosure variables based on sustainability reports. Due to limited availability, we have a lower number of observations. We use a word window of 14. The dependent variable is the ESG rating from MSCI KLD. All variables are explained in detail in Appendix B. The sample covers the years 2017 and 2018. For a regression analysis based on the MSCI word list with the ESG rating from MSCI as dependent variable, based on sustainability reports, controls, and fixed-effects, for example, the regression coefficient of *Dummies_MSCI*, is 0.247. We use robust standard errors. The symbols ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels, respectively; t-statistics are displayed in parentheses.

state, information is often unambiguous and depends on the specific sentence and the firm's environment (e.g., industry).¹⁹ Additionally, Plumlee et al. (2015) divide their measures into negative and positive subcategories which reduces the overall explanatory power of the models and neither the environmental strengths nor the environmental concerns are statistically significant.

Another limitation is the subjectivity related in creating dictionaries. Although we rely on existing studies and guidelines, subjectivity remains in the final word list creation. The regressions reveal a high constant that indicates the systematic error, reflecting weaknesses in our

¹⁹ Different topics can be associated with strengths or weaknesses by adding words like 'reduction' or 'increase'. However, we cannot per se label 'reduction' to a strength, as there are contexts where reduction signifies a weakness. For instance, 'corruption' is considered negative according to the word list from Loughran and McDonald (2011). However, firms might describe strategies to address corruption, e.g., fighting corruption, anti-corruption measures, referencing specific policies or guidelines. Overall, identifying reliable indicators for strengths and weaknesses in the ESG-context is difficult.

word lists. We have to be aware that we cannot capture every ESG topic formulation (see, e.g., Kotsantonis & Serafeim, 2019). Even rating agencies face challenges transforming textual information into suitable ESG metrics. In a random sample of 50 listed firms in Fortune 500, Kotsantonis and Serafeim (2019) find more than 20 metrics used by firms to report on employee health and safety issues. Therefore, we acknowledge some limitations in the model fit. Even if metrics are similar, there is no consensus on which metric best explains good ESG activities (Christensen et al., 2022).

Rating agencies are essential financial intermediaries and ESG ratings are widely used in practice. These ratings can account for strengths and weaknesses, and are assumed to address issues such as greenwashing and information verification, which is highly relevant for financial market efficiency. However, it remains questionable whether firms with higher ESG ratings are the more sustainable firms (e.g., Dremptic et al., 2020). In our study, we do not offer a solution to this problem but we show the consistency between several ESG ratings and firm disclosure. The ESG disclosure and the ESG rating of firms show only small variations over time, indicating our results primarily rely on between-firm variance. Our results indicate that ESG disclosure from annual reports can explain a substantial proportion of the ESG rating from MSCI KLD and other rating agencies. ESG rating agencies utilize various public information besides annual reports to create ESG ratings such as (social) media (e.g., Nicolas et al., 2024). This finding is relevant for market participants, who can use firms' annual reports as an approximation of several ESG ratings. Our findings enhance the understanding of the information value derived from two key information channels. Firms have further incentives to provide useful and accurate ESG disclosure, even when they are not rated by ESG rating agencies which surprisingly is the case for many firms listed on stock exchanges (Horn, 2023). Our examination of information consistency is a valuable contribution and provides a basis for further research. We find in parts strong differences in the explanatory power of different ESG ratings underscoring the discussion on ESG rating dispersion (Horn & Oehler, 2024; Oehler & Horn, 2025). If market participants base their decisions on ESG ratings, they can independently extract sustainability information from annual reports using them as a proxy for several ESG ratings. This approach increases their investment universe and reduces their reliance on ESG-rated firms. However, it does not resolve the problem of identifying the most sustainable firms.

5. Conclusion

This paper examines the consistency between ESG disclosure and ESG ratings from different ESG rating agencies. We answer the research question, how similar both information sources are. This research question is important because market participants use both sources for decision-making. We conduct the analysis with two word lists reflecting the rating agency perspective and the academic perspective. The results indicate that the academic word list, which reflects a comprehensive list of ESG-related words, can explain a substantial part of several ESG ratings. The explanatory power of the disclosure variables varies between ESG ratings, highlighting the issue of rating dispersion. We find the highest explanatory power for ESG ratings from Refinitiv and the lowest for ratings from ISS, indicating that ISS likely relies on other information besides the annual report or employs more sophisticated methods in its analysis. While this reinforces the importance of public ESG disclosure as an information channel, it also raises questions about the added value of ratings if a large part of the rating can already be derived from publicly available documents. However, the unexplained portion of the ESG ratings may reflect valuable inputs, including third-party data, or other information which is not found in annual reports. Our results underscore the dual role of ESG disclosure. It is not only a regulatory and reputational tool for firms, but also a core input for ESG evaluations. This highlights the need for transparency in rating methodologies and provides insights for regulatory authorities and market

Table 13
Panel regression based on sustainability reports – other ratings.

Dependent variable =	ISS		Vigeo Eiris		Sustainalytics		Refinitiv	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Dummies_MSCI</i>	0.002 (0.001)		0.044 (0.033)		0.079** (0.036)		0.019 (0.064)	
<i>Sum_academic</i>		0.002*** (0.001)		0.056*** (0.015)		0.076*** (0.016)		0.061* (0.032)
Constant	1.968*** (0.172)	1.859*** (0.170)	26.691*** (4.903)	23.625*** (4.671)	57.571*** (4.433)	54.695*** (4.406)	44.692*** (9.236)	40.534*** (9.231)
Industry and time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	398	398	373	373	357	357	390	390
Adj. R ²	0.204	0.251	0.334	0.370	0.217	0.289	0.282	0.302

Notes: This table shows the impact of each ESG pillar individually as well as the overall ESG disclosure variables based on sustainability reports. Due to limited availability, we have a lower number of observations. We use a word window of 14. The dependent variables are the ESG ratings from several ESG rating agencies. All variables are explained in detail in [Appendix B](#). The sample covers sustainability reports from 2002 to 2019. For a regression analysis based on the MSCI word list with the ESG rating from ISS as dependent variable, based on sustainability reports, controls, and fixed-effects, for example, the regression coefficient of *Dummies_MSCI*, is 0.002. We use robust standard errors clustered by firm. The symbols ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels, respectively; t-statistics are displayed in parentheses.

Table 14
Principal component analysis.

Academic word list Dependent variable =	<i>MSCI KLD</i>	<i>Refinitiv</i>	<i>Sustainalytics</i>	<i>Vigeo Eiris</i>	<i>ISS</i>
	(1)	(2)	(3)	(4)	(5)
Scores for component 1	0.314 (1.321)	1.522*** (0.275)	0.784*** (0.136)	0.803*** (0.139)	0.004 (0.007)
...	./.	./.	./.	./.	./.
Scores for component 157	0.980 (4.328)	0.615* (0.317)	0.179 (0.153)	0.325* (0.183)	0.008 (0.010)
Constant	15.660 (10.731)	45.229*** (2.934)	57.876*** (1.337)	25.396*** (1.432)	2.172*** (0.059)
Industry and time FE	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
Obs.	212	839	804	925	1319
Adj. R ²	0.706	0.677	0.713	0.623	0.436

Notes: This table shows the regression results using the principal components as independent variables. We include 157 principal components; we only show the first and the last one to ensure clarity. The dependent variable is the ESG rating from MSCI KLD, Refinitiv, Sustainalytics, Vigeo Eiris or ISS. All variables are explained in detail in [Appendix B](#). For a regression analysis including control variable data and fixed effects, with the ESG rating from Refinitiv as dependent variable, for example, the regression coefficient of the 157th component is 0.615. The symbols ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels, respectively; t-statistics are displayed in parentheses.

participants who rely on one or both information channels. Moreover, the results show incentives for firms to provide useful and accurate ESG disclosure.

Our analysis shows that the ESG disclosure from annual reports can serve as an approximation for several ESG ratings. If market participants base their decisions on ESG ratings, they can independently extract sustainability information from annual reports, using them as a proxy for several ESG ratings. This approach broadens their investment universe and reduces their reliance on ESG-rated firms. However, the challenge of identifying the most sustainable firms remains. Given that the MSCI sample spans only two years, our results for this rating should be interpreted as primarily cross-sectional in nature. The consistency of findings across other ratings with longer time spans supports the robustness of our conclusions.

The academic word list includes the keywords published from other studies (see [Appendix F](#)). For our textual analysis, we translated the keywords into German, and this approach can be applied easily to other languages. Alternatively, a multi-language corpus can be translated into a target language (see [Borms et al., 2021](#)). Depending on the language, it can be helpful to use stemming in the pre-processing of the data. Future research can use our provided word list to investigate country-specific differences and to improve the textual analysis with other methods, e.

g., artificial intelligence (see e.g., [Bochkay et al., 2023](#); [Kartal et al., 2024](#)) to make it more applicable for market participants.

Another important avenue for future research is to investigate the impact of regulatory changes on ESG disclosure practices. Specifically, future research could explore whether the reporting directive from 2017 improved ESG reporting, thereby increasing the consistency between ESG disclosures and ESG ratings. Alternatively, it is possible that increased disclosure requirements introduced noise or inconsistencies across firms, leading to a weaker explanatory relationship.

Finally, future research should investigate how changes in disclosure norms affect the explanatory power of ESG-related text for sustainability assessments since recent developments suggest that some firms are scaling back ESG-related language in their public communication (e.g., [Falchi et al., 2022](#)). This could affect the link between ESG disclosure and ESG ratings by reduced transparency of the disclosure or changes in the rating agencies methods.

Declaration of competing interest

The authors have nothing to declare. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix A. Sample construction

Selection criteria	Firms	Obs.
Start: Listed non-financial firms in Germany (2002–2019) without survivorship bias	592	6909
With control variables data	524 (68)	5314 (1595)
(1) With ESG rating from MSCI	118 (406)	212 (5102)
(2) With ESG rating from different agencies	197 (327)	1618 (3696)
<i>Rating overview:</i>	<i>Firms</i>	<i>Observations</i>
With Rating ISS	168	1319
With Rating Vigeo Eiris	128	925
With Rating Refinitiv	114	839
With Rating Sustainalytics	118	804

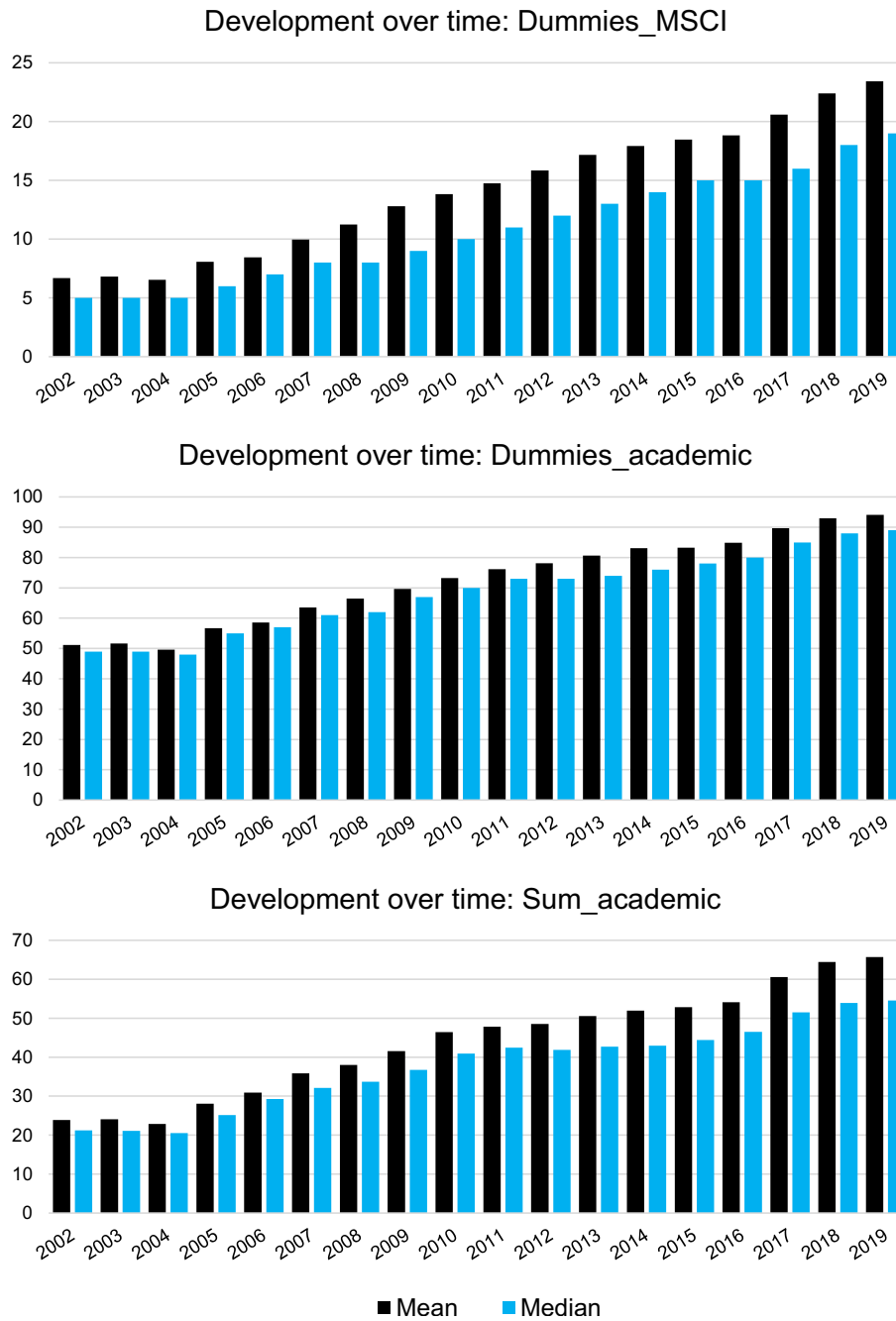
Note: The table presents the summary of our sample. The sample cover the period from 2002 to 2019 for German firms listed in CDAX. The final dataset shows firms that received an ESG Rating from MSCI KLD. Control variables are explained in [Appendix B](#). Example: From the first 592 firms, only 197 firms have at least one ESG rating from one of the mentioned ESG rating agencies. 118 firms have an ESG rating from MSCI KLD, 168 firms from ISS, 128 firms from Vigeo Eiris, 114 firms from Refinitiv, and 118 firms from Sustainalytics.

Appendix B. Variables definition

Variable name	Description	Source
ESG Variables		
<i>MSCI_Rating</i>	A firm’s ESG rating from MSCI KLD	MSCI KLD
<i>ISS</i>	A firm’s ESG rating from ISS	ISS
<i>Refinitiv</i>	A firm’s ESG rating from Refinitiv	Refinitiv
<i>Sustainalytics</i>	A firm’s ESG rating from Sustainalytics	Sustainalytics
<i>Vigeo Eiris</i>	A firm’s ESG rating from Vigeo Eiris	Vigeo Eiris
<i>Dummies_MSCI</i>	Firm’s ESG disclosure score from the annual report in year <i>t</i> . For each keyword we use a dummy variable (1 if reported) and the final measure is the sum of keywords (i.e., dummy variables) reported (between 0 and the 363); word window = 14 if no other window mentioned.	MSCI KLD ESG Rating Guideline (Textual analysis using Rapid Miner)
<i>Sum_MSCI</i>	Firm’s ESG disclosure score from the annual report in year <i>t</i> calculated as the absolute number of occurrences of keywords (includes repetitions of keywords) divided by the log of the annual reports’ length; word window = 14. The variable is winsorized at 1 % and 99 % levels.	
<i>Dummies_academic</i>	Firm’s ESG disclosure score from the annual report in year <i>t</i> . For each keyword we use a dummy variable (1 if reported) and the final measure is the sum of keywords (i.e., dummy variables) reported (between 0 and the 646); word window = 14.	Published word lists from 12 studies
<i>Sum_academic</i>	Firm’s ESG disclosure score from the annual report in year <i>t</i> calculated as the sum of all occurrences of keywords; word window = 14.	
<i>E_MSCI</i>	Firms’ environmental / social / governance disclosure score from the annual report in year <i>t</i> ; calculated as the sum of keywords reported.	MSCI KLD ESG Rating Guideline
<i>S_MSCI</i>	Note: Prefix <i>S_</i> indicates calculation as sum (absolute number of occurrences of keywords)	
<i>G_MSCI</i>		
<i>E_academic</i>	Firms’ environmental / social / governance disclosure score from the annual report in year <i>t</i> ; calculated as the sum of keywords reported.	Published word lists from 12 studies
<i>S_academic</i>	Note: Prefix <i>S_</i> indicates calculation as sum (absolute number of occurrences of keywords)	
<i>G_academic</i>	Subcategories measured as sum of single dummy variables, i.e. the number of keywords reported:	
<i>Dummies_pollution</i>	‘Pollution and waste’	
<i>Dummies_climate</i>	‘Climate change’	
<i>Dummies_env_opp</i>	‘Environmental opportunities’	MSCI KLD ESG Rating Guideline or allocation of the underlying study; the used word list is mentioned in the notes of the table
<i>Dummies_naturalcapital</i>	‘Natural capital’	
<i>Dummies_human</i>	‘Human capital’	
<i>Dummies_social_opp</i>	‘Social opportunities’	
<i>Dummies_product</i>	‘Product liabilities’	
<i>Dummies_stakeholder</i>	‘Stakeholder opposition’	
<i>Dummies_behaviour</i>	‘Corporate behavior’	
<i>Dummies_governance</i>	‘Corporate governance’	
<i>Dummies_controversial</i>	‘Controversial involvement’	
<i>Sum_pollution</i>	Subcategories measured as the absolute number of occurrences of keywords relative to the number of keywords in each subcategory (see Appendix C):	
<i>Sum_climate</i>	‘Pollution and waste’	
<i>Sum_env_opp</i>	‘Climate change’	
<i>Sum_naturalcapital</i>	‘Environmental opportunities’	MSCI KLD ESG Rating Guideline or allocation of the underlying study; the used word list is mentioned in the notes of the table
<i>Sum_human</i>	‘Natural capital’	
<i>Sum_social_opp</i>	‘Human capital’	
<i>Sum_product</i>	‘Social opportunities’	
<i>Sum_stakeholder</i>	‘Product liabilities’	
<i>Sum_behaviour</i>	‘Stakeholder opposition’	
<i>Sum_governance</i>	‘Corporate behavior’	
<i>Sum_controversial</i>	‘Corporate governance’	
<i>Control Variables</i>		
<i>Years_listed</i>	Number of years a firm is listed at the stock market	Refinitiv Datastream
<i>MtB</i>	The market value of equity divided by the book value of equity at year <i>t</i>	Refinitiv Datastream
<i>Leverage</i>	Total liabilities over total assets at year <i>t</i>	Refinitiv Datastream
<i>Industry Fixed Effects</i>	Industry subgroup classification using the first two digits of four-digit numeric code	Refinitiv Datastream

Note: This table defines all variables.

Appendix C. Development over time of disclosure variables



Note: These graphs display the development of the noted disclosure scores between 2002 and 2019; we display mean and median. The variables are based on the whole dataset, i.e., 592 firms (see Appendix A).

Appendix D. Summary statistics based on ESG rating from MSCI KLD

	Firm-year obs.					Differences		
	No.	Mean	SD	Min	Max	P50-P50	P25-P75	P10-P90
Sum_MSCI						7.55***	11.17***	15.66***
P10	15	7.42	3.31	2.30	13.97	(5.8141)	(5.6914)	(3.3971)
P25	53	7.67	3.67	2.30	18.74			

(continued on next page)

(continued)

	Firm-year obs.					Differences		
	No.	Mean	SD	Min	Max	P50-P50	P25-P75	P10-P90
<i>P50 below</i>	115	8.92	5.66	0.97	39.62			
<i>P50 above</i>	97	16.47	12.50	2.21	65.41			
<i>P75</i>	53	18.84	13.81	4.59	65.41			
<i>P90</i>	21	23.09	17.56	6.06	65.41			
Dummies_MSCI						16.87***	24.11***	27.39***
<i>P10</i>	15	25.8	10.32	6	42	(7.2716)	(7.8356)	(4.2592)
<i>P25</i>	53	23.43	9.75	6	51			
<i>P50 below</i>	115	26.77	12.97	5	61			
<i>P50 above</i>	97	43.64	20.48	11	97			
<i>P75</i>	53	47.54	20.17	18	97			
<i>P90</i>	21	53.19	23.25	21	97			
Sum_academic						32.08***	50.34***	59.72***
<i>P10</i>	15	62.65	20.80	22.08	88.57	(8.3451)	(11.6379)	(8.7429)
<i>P25</i>	53	65.80	21.68	22.08	123.81			
<i>P50 below</i>	115	73.62	27.06	14.02	136.39			
<i>P50 above</i>	97	105.70	28.84	33.78	137.32			
<i>P75</i>	53	116.14	22.84	56.26	137.32			
<i>P90</i>	21	122.37	19.78	78.47	137.32			
Dummies_academic						36.18***	51.36***	63.68***
<i>P10</i>	15	99.47	29.37	42	144	(7.7722)	(8.6027)	(5.1957)
<i>P25</i>	53	101.17	24.94	42	148			
<i>P50 below</i>	115	108.19	30.11	27	182			
<i>P50 above</i>	97	144.37	37.65	66	243			
<i>P75</i>	53	152.53	35.59	83	243			
<i>P90</i>	21	163.14	40.38	89	243			
Years_listed						7.76***	11***	12.54***
<i>P10</i>	15	16.27	8.10	1	29	(4.6232)	(4.7191)	(2.8004)
<i>P25</i>	53	14.77	7.97	1	30			
<i>P50 below</i>	115	16.82	9.78	1	45			
<i>P50 above</i>	97	24.58	14.52	2	45			
<i>P75</i>	53	25.77	14.98	3	45			
<i>P90</i>	21	28.81	15.51	4	45			
MtB						0.95***	1.50***	1.99**
<i>P10</i>	15	4.54	4.01	0.78	11.7	(2.6423)	(2.8217)	(2.1156)
<i>P25</i>	53	4.29	3.62	0.78	12.93			
<i>P50 below</i>	115	3.72	3.11	0.45	13.61			
<i>P50 above</i>	97	2.76	1.86	0.59	14.25			
<i>P75</i>	53	2.79	1.38	0.59	6.4			
<i>P90</i>	21	2.55	1.37	0.59	5.52			
Leverage						0.07***	0.09***	0.09
<i>P10</i>	15	0.52	0.15	0.25	0.74	(3.1004)	(2.9789)	(1.6892)
<i>P25</i>	53	0.50	0.17	0.16	0.82			
<i>P50 below</i>	115	0.51	0.17	0.09	0.90			
<i>P50 above</i>	97	0.58	0.16	0.14	0.92			
<i>P75</i>	53	0.59	0.15	0.34	0.92			
<i>P90</i>	21	0.61	0.16	0.37	0.88			

Notes: This table shows the summary statistics of the sample of 118 firms grouped by their ESG rating from MSCI KLD. P10 consists of firms with the lowest value of the ESG rating and are in the 10th percentile. We report the mean, minimum, maximum, and the absolute values of differences. Differences are analyzed with *t*-tests with Welch's (1947) formula. We control the summary statistics for the other rating agencies and find similar results. We do not detect differences in the disclosure measures. Example: The mean value of *Dummies_MSCI* is 25.8 with a standard deviation of 10.32, and with a range from 6 to 42. The differences between the firms with the lowest *Dummies_MSCI* that are in the 10th percentile and firms that are in the 90th percentile is 27.39 and significant at the 1 % level. The symbols ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels, respectively; *t*-statistics are displayed in parentheses.

Appendix E. Summary statistics based *Dummies_MSCI* and *Sum_academic*

	Firm-year observations					Differences		
	No.	Mean	SD	Min	Max	P50-P50	P25-P75	P10-P90
Panel A: Word list MSCI (Dummies)								
<i>MSCI_Rating</i>	212	19.33	19.59	7.14	72.14	14.38***	22.48***	36.71***
						(5.7289)	(6.6105)	(8.9823)
<i>Years_listed</i>	1618	18.67	12.69	1	46	6.49***	10.10	11.38***
						(10.6855)	(10.7530)	(7.6240)
<i>MtB</i>	1618	2.40	2.99	48.07	39.67	0.18	0.55***	0.42
						(1.2366)	(2.8055)	(1.4069)
<i>Leverage</i>	1618	0.58	0.19	0.08	1.81	0.04***	0.05***	0.06***
						(4.2708)	(4.4152)	(2.9037)
<i>Refinitiv</i>	839	51.91	21.95	3.88	94.13	20.89***	31.46***	39.03***
						(14.4440)	(14.9600)	(13.5332)
<i>ISS</i>	1319	2.10	0.56	1.06	3.95	0.21***	0.30***	0.38***
						(6.7598)	(6.8551)	(5.8951)

(continued on next page)

(continued)

	Firm-year observations					Differences		
	No.	Mean	SD	Min	Max	P50-P50	P25-P75	P10-P90
<i>Sustainalytics</i>	804	59.48	10.86	37.83	86.25	9.36*** (12.5586)	14.63*** (11.7921)	20.24*** (9.2468)
<i>Vigeo Eiris</i>	925	34.20	11.77	8	63	10.74*** (15.0176)	15.22*** (14.5325)	20.26*** (13.2352)
Panel B: Word list academic (Sum)								
<i>MSCI_Rating</i>	212	19.33	19.59	7.14	72.14	20.91*** (9.1787)	29.55*** (9.4649)	38.31*** (9.7131)
<i>Years_listed</i>	1618	18.67	12.69	1	46	5.76*** (9.4177)	9.86*** (10.8391)	12.13*** (8.3012)
<i>MtB</i>	1618	2.40	2.99	48.07	39.67	0.13 (0.9414)	0.08 (0.5209)	0.08 (0.2366)
<i>Leverage</i>	1618	0.58	0.19	0.08	1.81	0.06*** (6.6840)	0.09*** (7.2688)	0.07*** (3.5422)
<i>Refinitiv</i>	839	51.91	21.95	3.88	94.13	20.71*** (14.1863)	27.89*** (13.8919)	36.59*** (14.4755)
<i>ISS</i>	1319	2.10	0.56	1.06	3.95	0.21*** (6.7850)	0.28*** (7.0085)	0.44*** (7.7732)
<i>Sustainalytics</i>	804	59.48	10.86	37.83	86.25	8.53*** (11.3828)	10.66*** (9.6395)	16.32*** (8.8255)
<i>Vigeo Eiris</i>	925	34.20	11.77	8	63	10.14*** (13.5276)	12.76*** (13.3538)	15.38*** (10.0389)

Notes: This table shows summarized statistics of the sample. In Panel A (B), we divide the sample into percentiles based on the Dummies_MSCI (Sum_academic) variable. We use a word window of 14. P10 auf Panel A consists of firms with the lowest number of ESG-related keywords reported in the annual report and are in the 10th percentile. We report the mean, minimum, maximum, and the absolute values of differences. Differences are analyzed with T-tests with Welch's (1947) formula. Example: The mean value of *Refinitiv* is 51.91 with a standard deviation of 21.95, and with a range from 3.88 to 94.13. The differences between the firms with low ESG disclosure that are in Panel A and in the 10th percentile and firms with high disclosure that are in the 90th percentile is 39.03 and significant at the 1 % level. The symbols ***, **, and * denote statistical significance at the 1 %, 5 %, and 10 % levels, respectively; t-statistics are displayed in parentheses.

Appendix F. Word list academic (only words with hits)

Word	Pillar	Source
waste treatment, waste, climate change mitigation, wastewater treatment, wastewater management, agricultural research, adaptability, extinction, species & threatened, mountain, housing & affordable, biodiversity, biological diversity, soil quality, drought, efficient utilization, resource utilization, efficient water use, energy efficiency, energy infrastructure, energy intensity, energy technology, desalination, renewable energy, extreme weather event, wetland, fish stock, fishing, illegal fishing, river, fossil fuel, global energy mix, trade restriction, catastrophe, climate adaptation, climate-related hazard, climate protection measure, agriculture, food loss, food waste, material usage, marine area, marine resource, marine technology, marine pollution, post-harvest loss, sustainable building, sustainable tourism, food price, food waste, natural disaster, recycling rate, reduction climate impact, resource footprint, clean technology, protected area, lake, fossil fuel subsidy, freshwater ecosystem, greenhouse gas emission, overfishing, flood, environmental pollution, environmentally friendly technology, diversity loss, air pollution, water pollution, soil pollution, forest management, forest, water management, water catastrophe, water scarcity, water use efficiency, water stress, resilience, recycling technology, reuse, recycling, waste reduction, waste minimization, waste avoidance, water pollution, air pollution, domestic wastewater, industrial wastewater	E	SDG (2022)
adequate hygiene, workplace accident, poverty, poverty line, enlightenment, exploitation, basic protection, vocational qualification, vocational training, blood poisoning, chronic respiratory disease, labor law compliance, income growth, safe drinking water, abduction, development planning, development cooperation, nutritional need, malnutrition, refugee, torture, health services, healthcare personnel, health risk, fundamental freedom, primary education, good water quality, HIV, hunger, vaccine, vaccination, incarceration, investment promotion program, youth employment, critical infrastructure, peace, culture, cultural diversity, malaria, human rights, migration, migration policy, sustainable procurement, sustainable livelihood, public performance, public transportation, sanitation services, job creation, school education, debt service, self-determination, safe drinking water supply, settlement planning, slavery, slum, death, killing, tropical disease, tuberculosis, financial market monitoring, undernutrition, poisoning, contraception, enforced disappearance, migrant worker, medication access, food access, forced labor, forced marriage, information technology access, communication technology access, energy service provider access, electricity access, university education, specialized education, vocational education, high-quality infrastructure, reliable infrastructure, sustainable infrastructure, resilient infrastructure, illness, fatality, sexual violence, accessible transportation system, sustainable transportation system	S	SDG (2022)
sustainability reporting, bribery, bribe payment, sustainability awareness, women in leadership position, early warning, gender equality, corruption, rule of law, risk reduction, increase/growth Gross Domestic Product, counterterrorism, crime & counteraction	G	SDG (2022)
waste and hazardous materials management, water and wastewater management, air quality, ecological impacts, energy management	E	Aldridge and Martin (2022)
affordability, customer welfare, product quality and safety, employee health and safety, labor practices, data security	S	
inclusion, systemic risk management, critical incident, risk management, competitive behavior	G	
cleaner, waste, biofuel, biphenyls, printing, emit, warming, plane, fossil, zoning, hazardous, poison, toxic, green, groundwater, house holding, contamination, air, sustainable pesticide, resource, clean, pollutants, species, nitrogen, wildlife, greenhouse, environmental, stewardship, deforestation, atmosphere, wetlands, coal, agriculture, solar, freshwater, recycle, wilderness, wind	E	Baier et al. (2020)
Word	Pillar	Source
ethnic, ethnically, ethnicities, ethnicity, gift, labor, workplace, educating, occupational, staffing, mentoring, bisexual, citizen, discriminate, eicc, hire, expression, fairness, woman, freedom, peace, warranty, gay, childbirth, society, health, marriage, hiv, homosexual, immigration, inspection, children, courses, wage, defects, medicine, overtime, people, humanity, headcount, ms, nations, nonprofit, pandemic, philanthropic, privacy, race, sexual, standardization, drinking, conformance, un, bargaining, injury, female, unemployment, poverty, disability, harassment, csr, epidemic, bugs, sick, illness, teach, teacher, learning, lesbian, lgbt, minerals, religion, scholarships, transgender, vulnerable, veteran, welfare, dignity	S	Baier et al. (2020)
recruiting, payout, bribery, visits, appreciation, assessment, relations, grassroots, bonus, brother, charter, drug, spousal, honesty, invite, parents, engagement, grandchildren, recoupment, detection, erm, elect, nomination, ethic, skill, parachute, family, fasb, leadership, gaap,	G	

(continued on next page)

(continued)

Word	Pillar	Source
salary, audit, auditor, grandparent, background, webpage, hotline, inform, incentive, insider, inspector, integrity, interview, investor, communicate, conflict, conformity, liaison, motivation, nephews, nieces, nominate, objectivity, duly, pension, plurality, presentation, press, examination, perspectives, qualifications, sister, son, interlocks, bylaw, vacancy, stepparents, stepchildren, talent, talented, tests, death, review, oversee, ungc, crimes, remuneration, posting, proxies, embezzlement, vote, ballot, alcohol, tenure, relatives, announce, notice, attracts, conduct, align, evaluate, lobbies, proponent, retain, vest, rewards, quorum, coso, influence, refreshment, fairly, feedback, misconduct, compensate, skill, control, corruption, culture, award, rotation, retirement, transparency, severance, perquisites, website, appreciation, whistleblower, mail, approval, consent, siblings		
indirect economic impacts, transport in environment, effluents, waste, materials	E	Bouten et al. (2011)
training, education, compliance of product, marketing communications, employee satisfaction, products and services, labor relations, occupational health and safety, procurement practices, customer satisfaction, security practices	S	
equal opportunity, compliance (labor practice), compliance (human rights), indigenous rights, market presence, anti-competitive behavior	G	
mobility, ecology, oil, oil leak, animal testing	E	Borms et al. (2021)
inclusion, society, trade union, mass fire, depression, slavery, strike	S	
money laundering, gender neutral, court, patent infringement, ethics, justice, lobbying, top wage	G	
environmental innovations and technologies, employee training in environmental management, donations related to environment, certification environmental programs, water use efficiency, air emissions, environmental awards, waste management, energy efficiency, environmental initiatives, environmental committee, environmental performance, certification & environmental impact	E	Clarkson et al. (2008)
ISO14001, environmental goals, GRI, environmental organizations, environmental performance, assurance about environmental information, department for pollution control, executive compensation & environment	G	
carbon trading, energy management, climate change strategies, environmental regulations, environmental certification, water consumption, energy use, environmentally responsible products, recycling facility	E	Chan et al. (2014)
improvement of working conditions, community relationships, community service, local suppliers, employee satisfaction, traineeship, donations, sponsorships, disabled people, migrants, injury prevention	S	
awards, air emission, reduce & spills, sites, goals, targets, conservation & natural resources, waste disposal information, noise, odors, spills, environmental management system, environmental policies, water discharge information, environmental debt	E	Cormier and Magnan (2003)
life cycle information, incidents	S	
finances, risk provision, compliance status of facilities, orders to conform, corrective actions	G	
effluents, energy consumption, environmental impacts recycled	E	Gamerschlag et al. (2011)
compliance, corruption, community, equal opportunities, employee turnover, product responsibility, sanctions, collective agreements, freedom of association, occupational health, customer health, customer safety, employment, occupational safety, forced labor	S	
Word	Pillar	Source
carbon footprint, global warming, greenhouse gas, floods, climate management, climate risk	E	Kouloukoui et al. (2019)
legal risk, regulatory risk, reputation risk, competitive risk	G	
alternative fuels, reduction carbon footprint, reduction gases, reduce fuel consumption, reduce spills, reduce packaging, minimize waste use, minimize water use, use recyclable materials, recycle waste, recycle water, reduce greenhouse gas emissions, reduced consumption of resources, reuse materials	E	Papoutsi and Sodhi (2020)
codes of conduct, assess/evaluate suppliers, source locally, source responsibly, collaborate with suppliers, procure sustainably, anti-corruption, product lifecycle assessment, train employees	S	
improvements & environment, environmental certification, environmental objectives, consumption, discharge, follow-up of environmental objectives	E	Tagesson et al. (2009)
equal opportunities, conditions of employment, investment policy, values, education of employees, safety and effect of the product	S	
code of conduct, investor relations	G	
conservation of natural resources, awards for environmental protection, departments or offices for pollution control, water discharge information	E	Wiseman (1982)
compliance status of facilities	G	
biodiversity, emissions, energy, waste management, environment, renewable energy, renewable, releases, ghg, ghg protocol, ghg emissions, ISO 14000, climate change, climate, carbon, land use, sustainability, natural disasters, recycling, reduce energy consumption, cleanup, environmental audits, environmental standards, environmental accidents, environment, pollution control, pollution, waste, water, storm	E	Multiple studies
charity, sponsoring, child labor, commit to employees, community, community support activities, customer privacy, discrimination, diversity, donation, employment, employee, employee engagement, endowment, equality, forced and compulsory labor, foundation, freedom of association, collective bargaining, gender, health, safety, human rights, community relations, labor, Medicaid, medicare, minorities, nondiscrimination, physical risk, product and service labelling, selling practices, product labeling, sex, social, sponsoring, training, women, minorities	S	
business ethics, diversity, independence, compliance, governance	G	
Excluded words	Reason	
Financing, consequence, proposal, green, price, resource, location, target, discharge, raw material, connection	Relation to ESG is not necessarily given and keywords are frequently reported.	
security risk, disclosure, stakeholder, safe, tests, management, supply chain		

Note: This table displays the keywords used for this study. We group the keywords by the ESG pillars and by their original study. At the end of the table, we show keywords that are not considered in the textual analysis.

Appendix G. Distribution keywords

Word list	Pillar / subcategory	No. of Words	Proportion
Academic	E	228	35.3 %
Academic	S	241	37.3 %
Academic	G	177	27.4 %
MSCI KLD	E	127	35 %
MSCI KLD	S	190	52 %
MSCI KLD	G	46	13 %
Academic	S - human capital	160	24.8 %
Academic	G - corporate governance	111	17.2 %
Academic	E - climate change	90	13.9 %

(continued on next page)

(continued)

Word list	Pillar / subcategory	No. of Words	Proportion
Academic	E - natural capital	71	11.0 %
Academic	G - corporate behavior	65	10.1 %
Academic	E - pollution and waste	45	7.0 %
Academic	S - social opportunities	40	6.2 %
Academic	S - product liability	27	4.2 %
Academic	E - environmental opportunities	24	3.7 %
Academic	S - stakeholder opposition	11	1.7 %
Academic	G - controversial	2	0.3 %

Note: This table displays an overview of the distribution of keywords for the used word lists. Example: Based on the academic word list, the G-pillar contains 177 words, which reflect 27.4 % of all words of the academic word list.

Data availability

The authors do not have permission to share data.

References

- Abdi, H., & Williams, L. J. (2010). Principal component analysis. *WIREs. Computational Statistics*, 2(4), 433–459. <https://doi.org/10.1002/wics.101>
- Abhayawansa, S., & Tyagi, S. (2021). Sustainable investing: The black box of environmental, social and governance (ESG) ratings. *Journal of Wealth Management*, 24(1), 49–54. <https://doi.org/10.2139/ssrn.3777674>
- Albarrak, M. S., Elnahass, M., & Salama, A. (2019). The effect of carbon dissemination on cost of equity. *Business Strategy and the Environment*, 28(6), 1179–1198. <https://doi.org/10.1002/bse.2310>
- Aldridge, I., & Martin, P. (2022). ESG in corporate filings: An AI perspective. *SSRN Journal*. <https://doi.org/10.2139/ssrn.4279479>
- Amel-Zadeh, A., & Serafeim, G. (2018). Why and how investors use ESG information: Evidence from a global survey. *Financial Analysts Journal*, 74(3), 87–103. <https://doi.org/10.2469/faj.v74.n3.2>
- Aureli, S., Del Baldo, M., Lombardi, R., & Nappo, F. (2020). Nonfinancial reporting regulation and challenges in sustainability disclosure and corporate governance practices. *Business Strategy and the Environment*, 29(6), 2392–2403. <https://doi.org/10.1002/bse.2509>
- Avetisyan, E., & Hockerts, K. (2017). The consolidation of the ESG rating industry as an enactment of institutional retrogression. *Business Strategy and the Environment*, 26(3), 316–330. <https://doi.org/10.1002/bse.1919>
- Baier, P., Berninger, M., & Kiesel, F. (2020). Environmental, social and governance reporting in annual reports: A textual analysis. *Financial Markets, Institutions & Instruments*, 29(3), 93–118. <https://doi.org/10.1111/fmii.12132>
- Baldini, M., Maso, L. D., Liberatore, G., Mazzi, F., & Terzani, S. (2018). Role of country- and firm-level determinants in environmental, social, and governance disclosure. *Journal of Business Ethics*, 150(1), 79–98. <https://doi.org/10.1007/s10551-016-3139-1>
- Balluchi, F., Lazzini, A., & Torelli, R. (2021). Credibility of environmental issues in non-financial mandatory disclosure: Measurement and determinants. *Journal of Cleaner Production*, 288, Article 125744. <https://doi.org/10.1016/j.jclepro.2020.125744>
- Ben-Porath, E., Dekel, E., & Lipman, B. L. (2018). Disclosure and choice. *The Review of Economic Studies*, 85(3), 1471–1501. <https://doi.org/10.1093/restud/rdx064>
- Benuzzi, M., Klaser, K., & Bax, K. (2024). Which ESG+F dimension matters most to private investors? An experimental study on financial decisions and future generations. *Journal of Behavioral and Experimental Finance*, 41, Article 100882. <https://doi.org/10.1016/j.jbef.2023.100882>
- Berg, F., Heeb, F., & Kölbl, J. F. (2022). The economic impact of ESG ratings. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4088545>
- Berg, F., Kölbl, J. F., & Rigobon, R. (2022). Aggregate confusion: The divergence of ESG ratings. *Review of Finance*, 26(6), 1315–1344. <https://doi.org/10.1093/rof/rfac033>
- Bialkowski, J., & Starks, L. T. (2016). *SRI funds: Investor demand, exogenous shocks and ESG profiles*.
- Billio, M., Costola, M., Hristova, I., Latino, C., & Pelizzon, L. (2021). Inside ESG ratings: (dis)agreement and performance. *Corporate Social Responsibility and Environmental Management*, 28(5), 1426–1445. <https://doi.org/10.1002/csr.2177>
- Bird, R., D Hall, A., Momentè, F., & Reggiani, F. (2007). What corporate social responsibility activities are valued by the market? *Journal of Business Ethics*, 76(2), 189–206. <https://doi.org/10.1007/s10551-006-9268-1>
- Bochkay, K., Brown, S. V., Leone, A. J., & Tucker, J. W. (2023). Textual analysis in accounting: What's next? *Contemporary Accounting Research*, 40, 765–805. <https://doi.org/10.1111/1911-3846.12825>
- Bonacorsi, L., Cerasi, V., Paola, G., & Manera, M. (2024). ESG factors and firms' credit risk. *Journal of Climate Finance*, 6, Article 100032. <https://doi.org/10.1016/j.jclimf.2024.100032>
- Borms, S., Boudt, K., Holle, F. V., & Willems, J. (2021). Semi-supervised text mining for monitoring the news about the ESG performance of companies. In S. Consoli, D. Reforgiato Recupero, & M. Saisana (Eds.), *Data science for economics and finance* (pp. 217–239). Springer International Publishing. https://doi.org/10.1007/978-3-030-66891-4_10
- Bouten, L., Everaert, P., van Liedekerke, L., Moor, L., & Christiaens, J. (2011). Corporate social responsibility reporting: A comprehensive picture? *Accounting Forum*, 35(3), 187–204. <https://doi.org/10.1016/j.accfor.2011.06.007>
- Bro, R., & Smilde, A. K. (2014). Principal component analysis. *Analytical Methods*, 6(9), 2812–2831. <https://doi.org/10.1039/C3AY41907J>
- Brown, N. C., Crowley, R. M., & Elliott, W. B. (2020). What are you saying? Using topic to detect financial misreporting. *Journal of Accounting Research*, 58(1), 237–291. <https://doi.org/10.1111/1475-679X.12294>
- Buallay, A. (2019). Is sustainability reporting (ESG) associated with performance? Evidence from the European banking sector. *Management of Environmental Quality: An International Journal*, 30(1), 98–115. <https://doi.org/10.1108/MEQ-12-2017-0149>
- Chan, M.-C. C., Watson, J., & Woodliff, D. (2014). Corporate governance quality and CSR disclosures. *Journal of Business Ethics*, 125(1), 59–73. <https://doi.org/10.1007/s10551-013-1887-8>
- Chatterji, A. K., Durand, R., Levine, D. I., & Touboul, S. (2016). Do ratings of firms converge? Implications for managers, investors and strategy researchers. *Strategic Management Journal*, 37(8), 1597–1614. <https://doi.org/10.1002/smj.2407>
- Christensen, D., Serafeim, G., & Sikoichi, A. (2022). Why is corporate virtue in the eye of the beholder? The case of ESG ratings. *The Accounting Review*, 97(1), 147–175. <https://doi.org/10.2308/TAR-2019-0506>
- Christensen, Hail, L., & Leuz, C. (2021). Mandatory CSR and sustainability reporting: Economic analysis and literature review. *Review of Accounting Studies*, 26(3), 1176–1248. <https://doi.org/10.1007/s11142-021-09609-5>
- Clarkson, P. M., Fang, X., Li, Y., & Richardson, G. (2013). The relevance of environmental disclosures: Are such disclosures incrementally informative? *Journal of Accounting and Public Policy*, 32(5), 410–431. <https://doi.org/10.1016/j.jaccpubpol.2013.06.008>
- Clarkson, P. M., Li, Y., Richardson, G. D., & Vasvari, F. P. (2008). Revisiting the relation between environmental performance and environmental disclosure: An empirical analysis. *Accounting, Organizations and Society*, 33(4–5), 303–327. <https://doi.org/10.1016/j.aos.2007.05.003>
- Clementino, E., & Perkins, R. (2021). How do companies respond to environmental, social and governance (ESG) ratings? Evidence from Italy. *Journal of Business Ethics*, 171(2), 379–397. <https://doi.org/10.1007/s10551-020-04441-4>
- Coad, A., Holm, J. R., Krafft, J., & Quattraro, F. (2018). Firm age and performance. *Journal of Evolutionary Economics*, 28(1), 1–11. <https://doi.org/10.1007/s00191-017-0532-6>
- Cormier, D., & Magnan, M. (2003). Environmental reporting management: A continental European perspective. *Journal of Accounting and Public Policy*, 22(1), 43–62. [https://doi.org/10.1016/S0278-4254\(02\)00085-6](https://doi.org/10.1016/S0278-4254(02)00085-6)
- Dhaliwal, D. S., Li, O. Z., Tsang, A., & Yang, Y. G. (2011). Voluntary nonfinancial disclosure and the cost of equity capital: The initiation of corporate social responsibility reporting. *The Accounting Review*, 86(1), 59–100. <https://doi.org/10.2308/accr-00000005>
- Dhaliwal, D. S., Radhakrishnan, S., Tsang, A., & Yang, Y. G. (2012). Nonfinancial disclosure and analyst forecast accuracy: International evidence on corporate social responsibility disclosure. *The Accounting Review*, 87(3), 723–759. <https://doi.org/10.2308/accr-10218>
- Dimson, E., Marsh, P., & Staunton, M. (2020). Divergent ESG ratings. *The Journal of Portfolio Management*, 47(1), 75–87. <https://doi.org/10.3905/jpm.2020.1.175>
- Ding, W., Levine, R., Lin, C., & Xie, W. (2022). Competition laws, ownership, and corporate social responsibility. *Journal of International Business Studies*, 53(8), 1576–1602. <https://doi.org/10.1057/s41267-022-00536-4>
- Dorfleitner, G., Halbritter, G., & Nguyen, M. (2015). Measuring the level and risk of corporate responsibility – An empirical comparison of different ESG rating approaches. *Journal of Asset Management*, 16(7), 450–466. <https://doi.org/10.1057/jam.2015.31>
- Dremptic, S., Klein, C., & Zwergel, B. (2020). The influence of firm size on the ESG score: Corporate sustainability ratings under review. *Journal of Business Ethics*, 167(2), 333–360. <https://doi.org/10.1007/s10551-019-04164-1>
- Dyck, A., Lins, K. V., Roth, L., & Wagner, H. F. (2019). Do institutional investors drive corporate social responsibility? International evidence. *Journal of Financial Economics*, 131(3), 693–714. <https://doi.org/10.1016/j.jfineco.2018.08.013>
- Eggertsson, T. (2013). Quick guide to new institutional economics. *Journal of Comparative Economics*, 41(1), 1–5. <https://doi.org/10.1016/j.jce.2013.01.002>
- Eng, L. L., Fikru, M., & Vichitsarawong, T. (2022). Comparing the informativeness of sustainability disclosures versus ESG disclosure ratings. *Sustainability Accounting*,

- Management and Policy Journal*, 13(2), 494–518. <https://doi.org/10.1108/SAMPJ-03-2021-0095>
- Falchi, A., Grolleau, G., & Mzoughi, N. (2022). Why companies might under-communicate their efforts for sustainable development and what can be done? *Business Strategy and the Environment*, 31(5), 1938–1946. <https://doi.org/10.1002/bse.2991>
- Ferjančić, U., Ichev, R., Lončarski, I., Montariol, S., Pelicon, A., Pollak, S., ... Žnidarišič, M. (2024). Textual analysis of corporate sustainability reporting and corporate ESG scores. *International Review of Financial Analysis*, 103669. <https://doi.org/10.1016/j.irfa.2024.103669>
- Galema, R., & Gerritsen, D. (2025). ESG rating changes and stock returns. *Journal of International Money and Finance*, 154, 103309.
- Gallo, P. J., & Christensen, L. J. (2011). Firm size matters: An empirical investigation of organizational size and ownership on sustainability-related behaviors. *Business & Society*, 50(2), 315–349. <https://doi.org/10.1177/0007650311398784>
- Gamerschlag, R., Möller, K., & Verbeeten, F. (2011). Determinants of voluntary CSR disclosure: Empirical evidence from Germany. *Review of Managerial Science*, 5(2–3), 233–262. <https://doi.org/10.1007/s11846-010-0052-3>
- Giles, O., & Murphy, D. (2016). SLAPPed: The relationship between SLAPP suits and changed ESG reporting by firms. *Sustainability Accounting, Management and Policy Journal*, 7(1), 44–79. <https://doi.org/10.1108/SAMPJ-12-2014-0084>
- Gilje, E. P. (2016). Do firms engage in risk-shifting? Empirical evidence. *Review of Financial Studies*, 29(11), 2925–2954. <https://doi.org/10.1093/rfs/hhw059>
- Gillan, S. L., Koch, A., & Starks, L. T. (2021). Firms and social responsibility: A review of ESG and CSR research in corporate finance. *Journal of Corporate Finance*, 66, Article 101889. <https://doi.org/10.1016/j.jcorpfin.2021.101889>
- Glück, M., Hübel, B., & Scholz, H. (2022). ESG rating events and stock market reactions. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3803254>
- Guerrero, S., & Viteri, J. P. (2025). What are environmental, social, and governance scores measuring? The role of outcome and impact indicators in ESG scores. *Finance Research Letters*, 72, Article 106529. <https://doi.org/10.1016/j.frl.2024.106529>
- Hahn, R., & Kühnen, M. (2013). Determinants of sustainability reporting: A review of results, trends, theory, and opportunities in an expanding field of research. *Journal of Cleaner Production*, 59, 5–21. <https://doi.org/10.1016/j.jclepro.2013.07.005>
- Haniffa, R. M., & Cooke, T. E. (2005). The impact of culture and governance on corporate social reporting. *Journal of Accounting and Public Policy*, 24(5), 391–430. <https://doi.org/10.1016/j.jaccpubpol.2005.06.001>
- Hartzmark, S. M., & Sussman, A. B. (2019). Do Investors value sustainability? A natural experiment examining ranking and fund flows. *The Journal of Finance*, 74(6), 2789–2837. <https://doi.org/10.1111/jofi.12841>
- Heichl, V., & Hirsch, S. (2023). Sustainable fingerprint – Using textual analysis to detect how listed EU firms report about ESG topics. *Journal of Cleaner Production*, 426, Article 138960. <https://doi.org/10.1016/j.jclepro.2023.138960>
- Hooks, J., & van Staden, C. J. (2011). Evaluating environmental disclosures: The relationship between quality and extent measures. *The British Accounting Review*, 43(3), 200–213. <https://doi.org/10.1016/j.bar.2011.06.005>
- Horn, M. (2023). The influence of ESG ratings on idiosyncratic stock risk: The unrated, the good, the bad, and the sinners. *Schmalenbachs Journal of Business Research*, 1–28. <https://doi.org/10.1007/s41471-023-00155-1>
- Horn, M., & Oehler, A. (2024). Constructing stock portfolios by sorting on ESG ratings: Does the rating provider matter? *International Review of Financial Analysis*, 96, Article 103568. <https://doi.org/10.1016/j.irfa.2024.103568>
- Huang, S., Lu, H., & Zhang, Y. (2024). *Disclosure Regulation and ESG Rating Disagreement*. <https://doi.org/10.2139/ssrn.5073527>
- Hull, C. E., & Rothenberg, S. (2008). Firm performance: The interactions of corporate social performance with innovation and industry differentiation. *Strategic Management Journal*, 29(7), 781–789. <https://doi.org/10.1002/smj.675>
- Hurdle, G. J. (1974). Leverage, risk, market structure and profitability. *The Review of Economics and Statistics*, 56(4), 478. <https://doi.org/10.2307/1924463>
- Jackson, D. A. (1993). Stopping rules in principal components analysis: A comparison of Heuristical and statistical approaches. *Ecology*, 74(8), 2204–2214. <https://doi.org/10.2307/1939574>
- Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065), 20150202. <https://doi.org/10.1098/rsta.2015.0202>
- Kartal, M. T., Taşkın, D., Şahbaz, M., Kılıç Depren, S., & Korkut Pata, U. (2024). Effects of environment, social, and governance (ESG) disclosures on ESG scores: Investigating the role of corporate governance for publicly traded Turkish companies. *Journal of Environmental Management*, 368, Article 122205. <https://doi.org/10.1016/j.jenvman.2024.122205>
- Khan, M. J., Hussain, D., & Mehmood, W. (2016). Why do firms adopt enterprise risk management (ERM)? Empirical evidence from France. *Management Decision*, 54(8), 1886–1907. <https://doi.org/10.1108/MD-09-2015-0400>
- Khatiri, I., & Kjærland, F. (2023). Sustainability reporting practices and environmental performance amongst nordic listed firms. *Journal of Cleaner Production*, 418, Article 138172. <https://doi.org/10.1016/j.jclepro.2023.138172>
- Kim, K., Patro, S., & Pereira, R. (2017). Option incentives, leverage, and risk-taking. *Journal of Corporate Finance*, 43, 1–18. <https://doi.org/10.1016/j.jcorpfin.2016.12.003>
- Kimbrough, M. D., Wang, X., Wei, S., & Zhang, J. (2024). Does voluntary ESG reporting resolve disagreement among ESG rating agencies? *The European Accounting Review*, 33(1), 15–47. <https://doi.org/10.1080/09638180.2022.2088588>
- Kotsantonis, S., & Serafeim, G. (2019). Four things no one will tell you about ESG data. *Journal of Applied Corporate Finance*, 31(2), 50–58. <https://doi.org/10.1111/jacf.12346>
- Kouloukoui, D., Sant'Anna, A. M. O., da Silva Gomes, S. M., Oliveira Marinho, M. M., Jong, P., Kiperstok, A., & Torres, E. A. (2019). Factors influencing the level of environmental disclosures in sustainability reports: Case of climate risk disclosure by Brazilian companies. *Corporate Social Responsibility and Environmental Management*, 26(4), 791–804. <https://doi.org/10.1002/csr.1721>
- van der Laan Smith, J., Adhikari, A., & Tondkar, R. H. (2005). Exploring differences in social disclosures internationally: A stakeholder perspective. *Journal of Accounting and Public Policy*, 24(2), 123–151. <https://doi.org/10.1016/j.jaccpubpol.2004.12.007>
- Lang, M., & Stice-Lawrence, L. (2015). Textual analysis and international financial reporting: Large sample evidence. *Journal of Accounting and Economics*, 60(2–3), 110–135. <https://doi.org/10.1016/j.jacceco.2015.09.002>
- Liang, H., & Renneboog, L. (2017). On the foundation of corporate social responsibility. *The Journal of Finance*, 72(2), 853–910. <https://doi.org/10.1111/jofi.12487>
- Lins, K. V., Servaes, H., & Tamayo, A. (2017). Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis. *The Journal of Finance*, 72(4), 1785–1824. <https://doi.org/10.1111/jofi.12505>
- Liu, M. (2022). Quantitative ESG disclosure and divergence of ESG ratings. *Frontiers in Psychology*, 13, Article 936798. <https://doi.org/10.3389/fpsyg.2022.936798>
- Lokuwaduge, C. S. D. S., & Heenetigala, K. (2017). Integrating environmental, social and governance (ESG) disclosure for a sustainable development: An Australian study. *Business Strategy and the Environment*, 26(4), 438–450. <https://doi.org/10.1002/bse.1927>
- Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, 66(1), 35–65. <https://doi.org/10.1111/j.1540-6261.2010.01625.x>
- Loughran, T., & McDonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, 54(4), 1187–1230. <https://doi.org/10.1111/1475-679X.12123>
- Loughran, T., McDonald, B., & Yun, H. (2009). A wolf in sheep's clothing: The use of ethics-related terms in 10-K reports. *Journal of Business Ethics*, 89(S1), 39–49. <https://doi.org/10.1007/s10551-008-9910-1>
- Manita, R., Bruna, M. G., Dang, R., & Houanti, L. (2018). Board gender diversity and ESG disclosure: Evidence from the USA. *Journal of Applied Accounting Research*, 19(2), 206–224. <https://doi.org/10.1108/JAAR-01-2017-0024>
- Manning, C., & Schütze, H. (2000). *Foundations of statistical natural language processing* (2nd ed.). MIT Press.
- Martin, P. R., & Moser, D. V. (2016). Managers' green investment disclosures and investors' reaction. *Journal of Accounting and Economics*, 61(1), 239–254. <https://doi.org/10.1016/j.jacceco.2015.08.004>
- Mbanye, W., & Muchenje, L. (2022). Climate change exposure, risk management and corporate social responsibility: Cross-country evidence. *Journal of Multinational Financial Management*, 66, Article 100771.
- Mishra, S., & Modi, S. B. (2013). Positive and negative corporate social responsibility, financial leverage, and idiosyncratic risk. *Journal of Business Ethics*, 117(2), 431–448. <https://doi.org/10.1007/s10551-012-1526-9>
- Muck, M., & Schmidt, T. (2024). Comparing ESG score weighting approaches and stock performance differentiation. *Finance Research Letters*, 67, Article 105924. <https://doi.org/10.1016/j.frl.2024.105924>
- Neuendorf, K. A. (2002). *The content analysis guidebook*. Sage Publications.
- Nicolas, M. L. D., Desroziers, A., Caccioli, F., & Aste, T. (2024). ESG reputation risk matters: An event study based on social media data. *Finance Research Letters*, 59, Article 104712. <https://doi.org/10.1016/j.frl.2023.104712>
- Oehler, A., & Horn, M. (2025). Contemporaneous ESG ratings and idiosyncratic stock risk: Empirical evidence on measures of market consensus and dispersion. *International Review of Economics & Finance*, 104471. <https://doi.org/10.1016/j.iref.2025.104471>
- Papoutsis, A., & Sodhi, M. S. (2020). Does disclosure in sustainability reports indicate actual sustainability performance? *Journal of Cleaner Production*, 260, Article 121049. <https://doi.org/10.1016/j.jclepro.2020.121049>
- Pelster, M., Horn, M., & Oehler, A. (2024). Who cares about ESG? *Journal of Climate Finance*, 8, Article 100045. <https://doi.org/10.1016/j.jclimf.2024.100045>
- Plumlee, M., Brown, D., Hayes, R. M., & Marshall, R. S. (2015). Voluntary environmental disclosure quality and firm value: Further evidence. *Journal of Accounting and Public Policy*, 34(4), 336–361. <https://doi.org/10.1016/j.jaccpubpol.2015.04.004>
- Prasad, D. B. (2008). Content analysis: A method of social science research. In *Research methods for social work* (pp. 173–193). <http://www.css.ac.in/download/deviprasad/content%20analysis.%20a%20method%20of%20social%20science%20research.pdf>
- Raimo, N., Caragnano, A., Zito, M., Vitolla, F., & Mariani, M. (2021). Extending the benefits of ESG disclosure: The effect on the cost of debt financing. *Corporate Social Responsibility and Environmental Management*, 28(4), 1412–1421. <https://doi.org/10.1002/csr.2134>
- Rajesh, R., & Rajendran, C. (2020). Relating environmental, social, and governance scores and sustainability performances of firms: An empirical analysis. *Business Strategy and the Environment*, 29(3), 1247–1267. <https://doi.org/10.1002/bse.2429>
- Ramsey, J. B. (1969). Tests for specification errors in classical linear least-squares regression analysis. *Journal of the Royal Statistical Society: Series B: Methodological*, 31(2), 350–371. <https://doi.org/10.1111/j.2517-6161.1969.tb00796.x>
- Reverte, C. (2016). Corporate social responsibility disclosure and market valuation: Evidence from Spanish listed firms. *Review of Managerial Science*, 10(2), 411–435. <https://doi.org/10.1007/s11846-014-0151-7>
- Sahin, Ö., Bax, K., Czado, C., & Paterlini, S. (2022). Environmental, social, governance scores and the missing pillar—Why does missing information matter? *Corporate Social Responsibility and Environmental Management*, 29(5), 1782–1798. <https://doi.org/10.1002/csr.2326>

- Schadewitz, H., & Niskala, M. (2010). Communication via responsibility reporting and its effect on firm value in Finland. *Corporate Social Responsibility and Environmental Management*. <https://doi.org/10.1002/csr.234>
- Schiemann, F., & Tietmeyer, R. (2022). ESG controversies, ESG disclosure and analyst forecast accuracy. *International Review of Financial Analysis*, 84, Article 102373. <https://doi.org/10.1016/j.irfa.2022.102373>
- Schimanski, T., Reding, A., Reding, N., Bingler, J., Kraus, M., & Leippold, M. (2024). Bridging the gap in ESG measurement: Using NLP to quantify environmental, social, and governance communication. *Finance Research Letters*, 61, Article 104979. <https://doi.org/10.1016/j.frl.2024.104979>
- SDG. (2022). Global indicator framework for the sustainable development goals and targets of the 2030 agenda for sustainable development. <https://unstats.un.org/sdgs/indicators/Global%20Indicator%20Framework%20after%202023%20refinement%20Eng.pdf>.
- Serafeim, G., & Yoon, A. (2022). Stock price reactions to ESG news: The role of ESG ratings and disagreement. *Review of Accounting Studies*. <https://doi.org/10.1007/s11142-022-09675-3>
- Tagesson, T., Blank, V., Broberg, P., & Collin, S.-O. (2009). What explains the extent and content of social and environmental disclosures on corporate websites: A study of social and environmental reporting in Swedish listed corporations. *Corporate Social Responsibility and Environmental Management*, 16(6), 352–364. <https://doi.org/10.1002/csr.194>
- Tamimi, N., & Sebastianelli, R. (2017). Transparency among S&P 500 companies: An analysis of ESG disclosure scores. *Management Decision*, 55(8), 1660–1680. <https://doi.org/10.1108/MD-01-2017-0018>
- Tarnaud, A. C., & Zakriya, M. (2023). CSR disclosures and firm value: Disentangling the role of ESG rating providers. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4274053>
- Tetlok, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, 62(3), 1139–1168. <https://doi.org/10.1111/j.1540-6261.2007.01232.x>
- Threlfall, R., King, A., Shulman, J., & Bartels, W. (2020). *The time has come: The KPMG survey of sustainability reporting 2020*.
- Verbeeten, F. H. M., Gamerschlag, R., & Möller, K. (2016). Are CSR disclosures relevant for investors? Empirical evidence from Germany. *Management Decision*, 54(6), 1359–1382. <https://doi.org/10.1108/MD-08-2015-0345>
- Verrecchia, R. E. (1983). Discretionary disclosure. *Journal of Accounting and Economics*, 5, 179–194. [https://doi.org/10.1016/0165-4101\(83\)90011-3](https://doi.org/10.1016/0165-4101(83)90011-3)
- Welch, B. L. (1947). The generalisation of student's problems when several different population variances are involved. *Biometrika*, 34(1–2), 28–35. <https://doi.org/10.1093/biomet/34.1-2.28>
- Wiseman, J. (1982). An evaluation of environmental disclosures made in corporate annual reports. *Accounting, Organizations and Society*, 7(1), 53–63. [https://doi.org/10.1016/0361-3682\(82\)90025-3](https://doi.org/10.1016/0361-3682(82)90025-3)