PhD-FDEF-2025-001
Faculty of Law, Economics and Finance

DISSERTATION

Defense held on 26 February 2025 in Luxembourg

to obtain the degree of

DOCTEUR DE L'UNIVERSITÉ DU LUXEMBOURG EN SCIENCES DE GESTION

by

Melanie LUXEMBOURGER
Born on 30 December 1996 in Villers-Semeuse (France)

COMPREHENSIVE ESG DISCLOSURES: THE COLLECTIVE PATH TO CORPORATE RESPONSIBILITY

Dissertation defense committee:

Dr. Thomas Kaspereit, dissertation supervisor
Associate Professor, Université du Luxembourg

Dr. Anke Müßig, Chairman
Professor, Université du Luxembourg

Dr. Michael Halling,
Professor, Université du Luxembourg

Dr. Sofia Margarida Morais Lourenço,
Associate Professor, University of Lisbon

Dr. Regis Blazy,
Professor, Université de Strasbourg

Contents

Introduction

"If you write the problem down clearly, then the matter is half solved."

Kidlin's Law

Under the Environment, Social, and Governance (ESG) framework, the financial sector is undergoing significant transformations to address the challenges facing global society. With the growing evidence of climate change and the materialization of non-financial risks, sustainability issues have risen to the global political agenda. The ESG framework aims to integrate non-financial considerations—namely ESG risks and opportunities—into traditional financial analysis. This focus extends beyond the investment sector and influences all sectors of the economy, as any organization seeking access to capital markets is now required to enhance transparency and disclose key sustainability-related indicators. The importance of ESG disclosures in promoting transparency and accountability within investment decisions has become particularly critical as all individuals—whether through direct investments, mutual funds, or pension and retirement plans—are increasingly impacted by these developments and need to be properly informed.

In a market economy, where the actions of one entity can have wide-reaching effects on others, every organization's activities contribute to a larger, interconnected system where one's impact affects the entire ecosystem. The collective responsibility of all market participants to comprehensively disclose their ESG impacts becomes crucial in this context. Without full disclosure from all participants, it is impossible to accurately assess the true impact of economic activities. Achieving true sustainability, therefore, requires a collective effort toward transparency and accountability. Sustainable finance, which seeks to align financial systems with sustainable development by integrating ESG factors into decision-making processes, depends on such collective action. This is essential as it mobilizes capital toward projects and initiatives that contribute to long-term environmental and social well-being, ensuring that economic growth does not come at the expense of future generations. Without transparency and accountability from all market participants, the goal of sustainable finance remains out of reach.

However the reliability of disclosed information has become a significant concern (Amel-Zadeh and Serafeim, 2018; Diouf and Boiral, 2017; van Duuren et al., 2016). Various forms of misrepresentation exist: "greenwashing," which involves the overstatement of environmental achievements, and "brownwashing," which refers to the understatement of these achievements. Firms can exploit flexible standards in ESG reporting, often using vague, boilerplate language that obscures their true actions (Crilly et al., 2016). In extreme cases, such as the Volkswagen scandal, companies may engage in deceptive practices that not only contradict their stated ESG commitments but also involve illegal activities to maintain a façade of sustainability leadership (Siano et al., 2017). The motivations behind these practices vary; firms often resort to greenwashing in anticipation of greater stakeholder engagement, while brownwashing becomes more prevalent in deregulated environments where shareholder interests dominate (Kim and Lyon, 2015).

Given these challenges, the transition from voluntary to mandatory ESG disclosure represents a critical step in ensuring access to sustainable information, making transparency and sustainability reporting mandatory for all large financial market participants. These regulations aim to enhance transparency, comparability, and accountability by enforcing specific standards. Research shows that mandatory disclosure not only increases ESG information but also improves its quality by imposing stricter reporting frameworks (Krueger et al., 2021). Moreover, external scrutiny from NGOs and regulatory bodies has been shown to effectively reduce misrepresentation practices (Kim and Lyon, 2015). However, the success of these regulations depends on a robust regulatory framework, effective enforcement, and accurate ESG data reporting. As ESG reporting evolves, collaboration among academics, practitioners, and policymakers will be essential in ensuring these disclosures meaningfully reflect organizational sustainability and advance sustainable finance.

To address these complexities, this thesis provides a comprehensive examination of the regulatory and strategic dimensions of ESG disclosures. The first chapter of this thesis presents the harmonization challenges faced during the initial efforts to regulate ESG disclosures at the EU level and assesses, based on a comparative analysis of research findings and country-cultural dimensions, the role of cultural and economic contexts in shaping the interpretation and application of these regulatory mandates. Notably, the chapter sheds light on the importance of material and strategic forward-looking information in enhancing the quality of these disclosures, aligning with the forthcoming regulatory revisions. Based on an experimental analysis, the second chapter offers insights into how the time-horizon precision of ESG disclosures influences their perceived credibility among stakeholders, exploring the joint effects of precision and stakeholder experience. The third chapter investigates the relationship between sustainability-forward-looking disclosures and stock liquidity, using empirical analysis to uncover the financial implications of precise ESG reporting. The fourth chapter explores the potential of generative AI in democratizing sustainability assessments, highlighting the opportunities and challenges in broadening access to ESG evaluations. Following this introduction:

Chapter 1: *Harmonizing ESG Reporting in Europe: Comparing research findings on EU Directive 2014/95/EU implementation and cultural dimensions*

The first chapter sets the stage by examining the regulatory landscape in Europe, particularly the implementation of the EU Directive 2014/95/EU across various countries. It assesses how national legal frameworks and societal expectations shape the quality and extent of ESG reporting. The analysis reveals significant regional differences, highlighting the importance of cultural and economic contexts in understanding how regulatory mandates are interpreted and applied at the national level.

Chapter 2: *Precision and Perception in ESG Reporting: Analyzing the Credibility in*

ESG Disclosures

Building on the regulatory overview, the second chapter delves into the methodologies and perceptions surrounding ESG reporting. It investigates the role of time horizon precision in ESG disclosures and how these are perceived by different stakeholders. The chapter analyzes how varying levels of precision in ESG reporting impact the perceived credibility of these reports. Through experimental design and analysis, the chapter explores the joint effects of precision and stakeholder experience on the evaluation of ESG reports.

Chapter 3: *Sustainability-Forward-Looking Disclosure and Stock Liquidity: A European Investigation*

The third chapter focuses on the implications of precise ESG reporting on financial markets, particularly examining the relationship between sustainability-forward-looking disclosures (SFLD) and stock liquidity. Through empirical analysis, the chapter explores how the quality of forward-looking sustainability disclosures impacts stock market liquidity, offering insights into the financial materiality of ESG practices.

Chapter 4: *Democratizing sustainability assessment: Leveraging generative AI for assessing corporate sustainability*

In the fourth chapter, the discussion turns to the potential of emerging technologies, such as AI, to enhance the accessibility of sustainability assessments. This chapter explores the challenges and opportunities posed by these technologies in broadening the scope of ESG evaluations, making them more inclusive and accessible to a wider range of stakeholders.

By examining both the overarching regulatory frameworks and the specific mechanisms that influence disclosure quality and market impact, this research aims to provide actionable insights into how ESG reporting can be optimized to meet both regulatory demands and stakeholder expectations.

1 References

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 (cit. on p. 6).

Crilly, D., Hansen, M., & Zollo, M. (2016). The grammar of decoupling: A cognitive-linguistic perspective on firms' sustainability claims and stakeholders' interpretation. *Academy of Management Journal*, *59*(2), 705–729. https://doi.org/10.5465/amj.2015.0171 (cit. on p. 6).

Diouf, D., & Boiral, O. (2017). The quality of sustainability reports and impression management: A stakeholder perspective. *Accounting, Auditing & Accountability Journal*, *30*(3), 643–667. https://doi.org/10.1108/AAAJ-04-2015-2044 (cit. on p. 6).

Kim, E. H., & Lyon, T. P. (2015). Greenwash vs. brownwash: Exaggeration and undue modesty in corporate sustainability disclosure. *Organization Science*, *26*(3), 705–723. https://doi.org/10.1287/orsc.2014.0949 (cit. on p. 6).

Krueger, P., Sautner, Z., Tang, D. Y., & Zhong, R. (2021). The effects of mandatory esg disclosure around the world. *Journal of Accounting Research*, *n/a*(n/a). https://doi.org/10.1111/1475-679X.12548 (cit. on p. 6).

Siano, A., Vollero, A., Conte, F., & Amabile, S. (2017). 'more than words': Expanding the taxonomy of greenwashing after the volkswagen scandal. *Journal of Business Research*, *71*, 27–37. https://doi.org/10.1016/j.jbusres.2016.11.002 (cit. on p. 6).

van Duuren, E., Plantinga, A., & Scholtens, B. (2016). Esg integration and the investment management process: Fundamental investing reinvented. *Journal of Business Ethics*, *138*(3), 525–533. https://doi.org/10.1007/s10551-015-2610-8 (cit. on p. 6).

Statement of Conjoint Work

The work presented in this dissertation includes both my independent research and joint research efforts with others.

Chapter Contributions

- **Chapter 1**: The research and writing for this chapter were conducted independently by me.
- **Chapter 2**: This project was a collaborative effort with Maxime Pettinger. Melanie Luxembourger designed the research project and conducted the survey. Both authors contributed equally to the analysis and the interpretation of the results.
- **Chapter 3**: This project was a collaborative effort with Imen Derouiche and Anke Müßig. Imen Derouiche designed the research project. Melanie Luxembourger collected the data. Both authors contributed to the analysis and the interpretation of the results.
- **Chapter 4**: This project was a collaborative effort with Maxime Pettinger. Melanie Luxembourger designed the study, collected the data, developed the automated evaluation processes, and interpreted the results. Both authors contributed equally to the analysis and the interpretation of the results.

I declare that, except where explicit reference is made to the contribution of others, this dissertation is the result of my independent work and has not been submitted for any other degree or professional qualification.

Chapter 1

Harmonizing ESG Reporting in Europe: Comparing research findings on EU Directive 2014/95/EU implementation and cultural dimensions

Melanie Luxembourger

University of Luxembourg, Luxembourg

Abstract

This study explores the role of country-specific legal frameworks and societal expectations on ESG reporting across Europe, particularly in the context of the EU Directive 2014/95/EU. The aim is to analyze how national cultural and economic contexts influence the interpretation, implementation, and effectiveness of these regulatory mandates, with a focus on identifying regional disparities in ESG reporting practices. The research synthesizes data from various academic studies on the implementation of the EU Directive 2014/95/EU across different European countries and identifies and categorizes the diverse national approaches to ESG reporting, drawing on cultural dimensions and economic conditions as key variables influencing the outcomes of the directive's implementation. The findings reveal significant disparities in the quality and extent of ESG reporting across Europe, driven largely by differences in national cultural values and economic contexts. Countries with strong pre-existing ESG practices, such as France and Scandinavian nations, show higher compliance and quality in reporting, whereas Central and Eastern European countries face challenges due to economic constraints and weaker governance structures. The research also highlights the role of cultural dimensions, such as power distance and individualism, in shaping the effectiveness of ESG disclosures. These results contribute to the ongoing discourse on ESG reporting by providing an understanding of how legal and cultural frameworks intersect to influence corporate sustainability practices in Europe.

1 Introduction

By implementing Directive 2014/95/EU, the European Union has made a significant stride towards corporate responsibility (Pizzi et al., 2021). This initiative is part of a larger European push to lead global sustainability standards and ensure that companies commit to and report on sustainable and responsible business operations. The Non-Financial Reporting Directive (Directive 2014/95/EU, hereafter NFRD), which came into effect in 2014, requires large companies with more than 500 employees to disclose non-financial and diversity information (European Commission, 2014). The directive mandates the inclusion of information on environmental matters, social and employee aspects, respect for human rights, and anti-corruption and bribery issues (Moggi et al., 2023). The directive aims to enhance corporate transparency and encourage sustainable practices by providing stakeholders with crucial non-financial information, thereby bridging informational gaps between companies and stakeholders (Grewal et al., 2019; Pizzi et al., 2022). About 11.000 companies are covered by the NFRD.

Research consistently indicates that the NFRD has significantly increased the quantity of non-financial disclosures (Agostini et al., 2022; Arif et al., 2022; Cordazzo et al., 2020; Korca et al., 2021; Papa et al., 2022), although the relationship with the quality of these disclosures is less significant (Agostini et al., 2022; Aureli et al., 2020; Bini et al., 2023), with significant enhancements observed in specific areas like environmental reporting (Papa et al., 2022), social and employee matters (Korca et al., 2021), policies and due diligence, principal risk, and non-financial key performance indicators (Traxler et al., 2023), while there is a persistence of low-quality stakeholder engagement disclosures (Petruzzelli and Badia, 2024) and business model reporting practices (Traxler et al., 2023). The directive has also positively influenced ESG performance, particularly in environmental and social aspects (Aluchna et al., 2022; Cuomo et al., 2024).

The NFRD is characterized by its flexibility, allowing Member States considerable discretion in its implementation (Szabados, 2021). National differences and institutional contexts have resulted in inconsistent application and effectiveness of the directive across the EU (Aureli et al., 2018; Dumitru et al., 2019; Posadas et al., 2023). Indeed, the non-prescriptive nature of the directive has led to varied interpretations and implementations, which has affected the overall quality and comparability of non-financial reports (Aureli et al., 2020; Pizzi et al., 2022; Posadas et al., 2023).

The heterogeneity of the real effects of the NFRD can be attributed to differences in country-level enforcement and implementation, highlighting the need for more research into these country-specific enforcement differences due to the limited information available on enforcement activities (Fiechter et al., 2022). As Gray, 1988 highlights, national accounting systems are influenced not only by economic and legal factors but also by cultural dimensions that shape how regulations are interpreted and enforced. The complexity of enforcing a unified reporting standard across diverse legal and cultural landscapes within the EU is further compounded by the important interrelationship between regulatory bodies and the accounting profession in various countries, as they form a 'regulatory community' where members 'frequently influence each other, act with reference to each other, and desire each other's respect' (Meidinger, 1987). This interdependence can lead to a form of 'cultural capture,' where the close relationship between these actors may result in biased enforcement or interpretation of the NFRD, exacerbating the inconsistencies observed across different Member States. Thus, this review explores the intersection of legal frameworks, societal norms, and cultural values in shaping ESG reporting practices across Europe.

Previous literature reviews on the Directive have identified several key gaps. Dinh et al., 2023 and Michalak et al., 2023 both discover a country-selection bias in the literature, with most studies focused on the UK, Germany, France, and Italy, and limited research on Eastern Europe. Initial studies on the NFRD originated from Denmark and Sweden, later expanding to other countries, with Italy becoming particularly prominent (Michalak et al., 2023). This geographical focus suggests that enforcement and academic scrutiny of the NFRD vary across different jurisdictions, with certain countries leading in both research output and policy development. Dinh et al., 2023 call for more research on the mandatory European sustainability reporting settings, including the transposition of

EU legislation into domestic law and the relevance of national enforcement. Furthermore, Korca et al., 2021 call for an examination of contextual factors in mandatory non-financial disclosures, the Directive's binding requirements and non-binding guidelines, and the Directive's effects across countries, organizations, and timelines.

In an attempt to investigate these gaps, the proposed review identifies relevant papers on how European nations implement the European Directive into their national law and the impact it had on their reporting practices. The study attempts to synthesis these data to understand how country-specific regulatory frameworks and social expectations affect ESG reporting in Europe, notably in the context of the EU Directive 2014/95/EU and answer the question of how do legal frameworks, societal norms, and cultural values affect ESG reporting methods in Europe?

To identify relevant academic articles for this review, a systematic search was conducted in databases like Scopus, Web of Science, EBSCOhost, Google Scholar, and the Social Science Research Network (SSRN), using a comprehensive query designed to capture various aspects of ESG reporting practices, specifically within the context of European countries and the impact of legal frameworks such as the NFRD and CSRD. The search terms included a combination of keywords related to ESG, sustainability, corporate social responsibility, and non-financial reporting, as well as specific country names within Europe to ensure geographical relevance.

The search query was structured as follows:

- Keywords for ESG and Reporting: Terms such as "Environmental, Social, and Governance," "ESG," "sustainab*," "Corporate social responsibility," "CSR," and "Non-financial" were paired with terms like "report*" and "disclos*" to capture a wide range of relevant literature on ESG reporting practices.
- Geographical Scope: The query included all European countries by name, as well as broader regional terms like "European Union" and "European Economic Area," ensuring the inclusion of studies from across the continent.

- Legal and Regulatory Frameworks: To focus on the regulatory context, terms such as "Non-Financial Reporting Directive," "NFRD," "CSRD," "Corporate Sustainability Reporting Directive," and combinations of "legal" or "regulat*" with "framework*" or "requirement*" were included.

Following the initial identification of studies through the systematic search, 522 records were retrieved from the selected databases. After removing 194 duplicates, 328 unique records were screened based on relevance to ESG reporting within the European context, particularly concerning the NFRD and CSRD. This led to the exclusion of 160 studies and the inability to retrieve 2, leaving 166 studies for full-text review. This phase aimed to confirm the relevance of each study to the research questions, resulting in the exclusion of 39 studies. Among these, 20 studies were excluded due to focusing on the wrong setting (e.g., non-European countries or regions), and 17 were excluded because they utilized study designs not aligned with the review's objectives. After applying all inclusion and exclusion criteria, 129 studies were deemed relevant and included in the final review. The detailed PRISMA flow diagram outlining this process is available in the Appendix A.

In line with Hardies et al., 2023 recommendations, this review does not restrict itself to a narrow set of 'leading' or 'top' journals, as this could exclude many relevant studies from less internationally recognized journals. This approach was particularly crucial for obtaining insights from specific countries, especially those that have been identified as understudied. By including respected regional publications[1], this review ensures that our analysis reflects both dominant trends and country-specific insights.

The remainder of the review is organized as follows. Section 2 reviews the different national implementations; it examines how the non-prescriptive nature of the directive has led to varied interpretations and implementations across EU countries. The analysis identifies relevant clusters of countries based on regional and cultural differences and discusses how these variations affect the overall quality and comparability of non-financial

[1]Such cases include for example studies like 'Development of Non-financial Reporting: The Case of Baltic Countries' published in European Integration Studies, a journal based in Lithuania, and 'Environmental Disclosures According to ESRS in Polish Companies' from the journal Economics and Environment, published by the Polish Association of Environmental and Resource Economists.

reports, highlighting the influence of national culture on ESG reporting practices. Section 3 explores the firm-specific cultural values and corporate governance structures that shape sustainability disclosures. Section 4 reviews the best practices according to EU non-binding guidelines and preparation for CSRD, focusing on materiality, business model, and strategy in the context of ESG reporting. It addresses the challenges of implementing these concepts and how the CSRD aims to address these issues. The need for forward-looking, strategic reporting is emphasized, along with the impact of cultural differences on the quality of forward-looking disclosures.

2 Country-level cultural dynamics: national implementations of the directive

This section reviews the different national implementations investigations in the literature. it examines how the non-prescriptive nature of the directive has led to varied interpretations and implementations across EU countries. To better understand these variations, an analysis of the frequency of particular European countries being the focus of research in the literature, whether alone or together with other countries.

[Insert Figure 1.1 about here]

This analysis reveals distinct clusters of countries that are frequently studied together, reflecting interest in their regional and cultural similarities in their approach to implementing the NFRD. Notable, Western and Southern European countries, Scandinavian countries, and Central and Eastern European (CEE) countries each form identifiable clusters.

2.1 Western Europe: strong pre-existing practices

Italy is the most extensively studied country, with 29 instances of research focus. A potential explanation is the fact that Italy is a nation with historically weak focus on sustainability (Posadas et al., 2023; Venturelli et al., 2019), where companies, on average, only fulfilled half of the NFRD requirements in their 2015 reports—just after the Directive was published but before it was enforced as national law (Venturelli et al., 2017). This context makes Italy a prime example for evaluating the effectiveness of regulatory enforcement, particularly given the country's exemplary implementation of the directive, especially regarding the necessary assurance (Aureli et al., 2018; Lombardi, Cosentino, et al., 2022). While numerous studies found a considerable increase in the quantity of disclosures post-implementation (Agostini et al., 2022; Papa et al., 2022), there is no evidence that regulatory pressures have positively impacted the quality of sustainability information disclosed (Posadas et al., 2023), with only a marginal improvement observed (Korca et al., 2021), and lack of in-depth presentation of information (Lombardi, Schimperna, et al., 2022). Firms with prior experience in non-financial reporting showed better compliance with the new directive but did not substantially change their practices to exceed the minimum requirements (Cordazzo et al., 2020; Doni et al., 2020). Several studies even found a significant decrease in the publication of sustainability performance indicators by Italian corporations (Raucci and Tarquinio, 2020; Tarquinio et al., 2020), suggesting a more selective and cautious approach to reporting, focusing on indicators deemed more relevant to comply with the Directive's requirements (Pizzi, 2024; Posadas et al., 2023).

While having similar legal frameworks and belonging to the same territorial context of Mediterranean countries (Posadas et al., 2023), Spanish culture is described as more indulgent compared to the more masculine, individualistic, and long-term oriented Italian culture. These cultural differences contribute to varying approaches and effectiveness in non-financial reporting between the two countries (García-Sánchez et al., 2023). Both countries indeed differ in sustainability reporting experience, as Spain had a stronger tradition in sustainability disclosure, promoting ESG practices through initiatives like CERSE in 2008 and the Spanish Sustainable Economic Law in 2011 (García-Sánchez et al., 2023; Posadas et al., 2023).

Similar to Italy, Spain made assurance of the reports mandatory, but also included a gradual increase in the scope from a 500-employee threshold at first to 250 from 2021 onwards, providing smaller companies with more time to prepare for the new requirements (Aguado-Correa et al., 2023; Anguiano-Santos and Rodriguez-Entrena, 2024). However, despite the mandate, non-compliance with some requirement is observed (García-Benau et al., 2022), notably with external assurance (Araya and Sierra-García, 2021). Studies collectively show that Spain's pre-existing culture and practices in sustainability reporting played a more crucial role in report quality, with voluntary and normative mechanisms proving more effective than coercive ones (Anguiano-Santos and Rodriguez-Entrena, 2024; Posadas et al., 2023). Although the NFRD increased the quantity, the extent of regulatory compliance was strongly linked to the specific business sector in which the company operates (Sierra-Garcia et al., 2018), with disclosure quality driven by a focus on topics relevant to influential stakeholders (Esteban-Arrea and Garcia-Torea, 2022).

The country with the strongest background in ESG reporting however is France, starting in 1977 with the "bilan social" requiring companies with over 300 employees to disclose standardized information on working conditions and remuneration and making it one of the first EU countries to mandate such practices (Aureli et al., 2018; Kühn et al., 2014). Followed the "Nouvelles Régulations Économiques" (NRE) Act of 2001, requiring the integration of social and environmental information into the annual reports of companies listed on the Paris stock exchange (Aureli et al., 2020; Baalouch et al., 2019; Kühn et al., 2014). Subsequent legislation, including the Grenelle I and II Acts (2009, 2010, 2012), expanded these obligations to large companies and subsidiaries of listed companies. The Grenelle II Act, effective from December 31, 2013, required companies with over 500 employees to produce annual "social and environmental reports," and necessitating third-party verification for credibility and transparency (Aureli et al., 2020; Baalouch et al., 2019; Kaya, 2016). The implementation of this comprehensive legislation has resulted in a higher level of compliance, as highlighted by the "ORÉE" 2014 study, which found that listed companies exhibited more systematic and accurate disclosures, to the extent that by 2013 nearly 100% of large companies were reporting non-financial information (Kaya, 2016). As such it is no surprise that France adopted the most stringent transposition of the Directive, notably expanding the scope of companies required to comply to non-listed companies and non-listed investment funds having a net turnover over EUR 100 million (Aureli et al., 2018; Heichl and Hirsch, 2023).

As two of the largest economies in the EU, France and Germany were often studied together, offering insights into how economically significant nations implement and influence ESG policies. While Germany did not mandate sustainability reporting prior to the Directive, the disclosure of nonfinancial information was already high, with 87% of the companies included in the DAX 30 disclosing nonfinancial information in 2010, and only two companies not doing it by 2016 (Mion and Adaui, 2019). However, driven by political and business interests, Germany was a fierce opponent of stricter ESG regulations and reporting rules, nearly causing the NFRD's failure (Raith, 2023). The German Parliament (Bundestag) eventually passed the CSR-RL-UG law, following a one-on-one principle without major changes despite opposition from Green and Left-Wing opposition parties (Gerwing et al., 2022; Raith, 2023). The implementation led to moderate changes in the quality between 2016 and 2017 (Mion and Adaui, 2019), but with significant positive development over the following years (Schröder, 2022). Moreover, while auditors must verify only the existence, not the content, of the nonfinancial statement according to § 317 German GAAP, most firms (51.8%) used external assurance for their nonfinancial statements, with the majority (93.9%) opting for limited assurance (Gerwing et al., 2022).

Similar to France, the UK had already mandated ESG disclosures under the Companies Act 2006 Regulations 2013, with several studies indicating a significant improvement of the quality and extent of sustainability-related disclosures among UK firms (Hamed et al., 2022; Hummel and Rotzel, 2019; Venturelli et al., 2019). However, in the transposition of the directive, the UK adopted a minimalistic approach, largely copying the directive's text without adding additional specifications or requirements (Aureli et al., 2018, 2020). Aureli et al., 2020 hypothesize that Brexit has probably contributed to this, as the UK government might not have had the motivation to add new requirements and increase regulatory uncertainty for businesses at the time. Moreover the UK can be viewed as a distinct case due to its unique 'liberal' market economy, which is marked by equity financing, dispersed ownership, and dynamic markets for corporate control, contrasting with the coordinated market economies common in Continental Europe (Aureli et al., 2018). However, the introduction of the Directive still significantly improved the quality of CSR reporting by firms (Hamed et al., 2022; Hummel and Rotzel, 2019; Venturelli et al., 2019). However, following its exit from the EU in 2020, the UK will no longer be required to adopt the upcoming CSRD. Instead, it will need to develop its own legislation to pursue climate neutrality (Lin, 2022).

Portugal and Greece were the only countries studied alone. Some explanation could be from their different implementation timelines, which could result in different stages of compliance and reporting practices, making it difficult to compare directly with other countries that implemented the directive on time. Greece incorporated the directive one year earlier in their national law than other countries (Breijer and Orij, 2022; Cuomo et al., 2024), while Portugal missed the original 2016 deadline for transposing the Directive into national law and was only transposed through the Decree-Law 89/2017, issued on July 28, 2017 (Carmo and Ribeiro, 2022). Moreover, Dias et al., 2017 stated that Portugal, Greece, and Spain were most affected by the financial crisis and the effects were pronounced and still ongoing at the time of their study in 2017, which could have influenced their corporate governance and reporting practices. Still, major effects of the Directive on the quality of non-financial disclosures were observed in the first year of implementation (Carmo and Ribeiro, 2022), with a clear increase in quality of disclosures for companies falling under the scope of NFRD (Lemos et al., 2023). However, Eugénio et al., 2022 identify that non-financial reporting and its assurance are still underdeveloped in Portugal, resulting in limited client demand and auditor experience in this area. This underdevelopment presents a significant challenge as the CSRD approaches.

2.2 Scandinavian countries: high cultural alignment

Scandinavian countries, specifically Denmark, Sweden, and Finland, are recognized for their advanced socio-economic and cultural environments, which foster robust ESG practices. These nations, known for their high quality of life, strong social dialogue, and effective social well-being policies, consistently lead in sustainable competitiveness (Krasodomska et al., 2022). Classified as "Countries with Well-Developed ESG Frameworks," their commitment to sustainable development and business ethics is influenced by cultural factors and regulatory environments that emphasize transparency and accountability (Pizzi et al., 2022; Singhania and Saini, 2022).

Scandinavian countries exhibit cultural traits of low power distance, high individualism and low masculinity, creating a conducive environment for transparent and ethical business practices, fostering egalitarianism, self-reliance, and a preference for cooperation and quality of life over competition (Habek and Wolniak, 2016; Pizzi et al., 2022; Zanellato and Tiron-Tudor, 2022). These cultural dimensions necessitate organizations to align their behaviors with societal expectations, fostering legitimacy and effective communication of ESG efforts (Khatri and Kjærland, 2023). The high levels of Sustainable Development Goals (SDGs) reporting in Sweden and Denmark reinforces the idea of a distinct "Scandinavian approach" to socially responsible practices and the significant role cultural dimensions play in socially responsible practices (Pizzi et al., 2022). Krasodomska et al., 2023 further finds that companies operating in countries with a higher level of SDGs achievement have strong social policies and cultural-cognitive pressures that encourage companies to engage in sustainability reporting and voluntary assurance.

Notably, Sweden and Denmark were part of the few countries that already integrated mandatory non-financial reporting into their legal systems before the EU Directive 2014/95/EU (Černe et al., 2017). Sweden mandated state-owned companies to publish annual sustainability reports according to Global Reporting Initiative (GRI) guidelines since 2008, while Denmark required large companies to report on ESG disclosures since 2009 (Khatri and Kjærland, 2023; Pizzi et al., 2022). Moreover, Denmark was the first country to transpose Directive 2014/95/EU into its national law, on May 21, 2015, applying it from 2016 to the largest listed and state-owned companies, and from 2018 to all large companies (Černe et al., 2017), which is not surprising since it largely because the Directive closely mirrored Denmark's existing domestic regulations (Kinderman, 2020).

They also both set the threshold at firms with more than 250 employees, lower than the EU directive's threshold of 500 employees (Andersson and Arvidsson, 2023; Černe et al., 2017). Finland expressed cautious support for the Directive, endorsing transparency while voicing concerns about potential administrative burdens and the possible adverse effects on cost efficiency and competitiveness (Kinderman, 2020).

One significant challenge that has been raised is the inherent cultural modesty prevalent in countries like Denmark (Lueg et al., 2016). This cultural trait often leads to under-communication of ESG achievements, as companies tend to downplay their successes to avoid seeming arrogant. This understated communication style can hinder the effectiveness of ESG initiatives by failing to fully engage stakeholders or adequately showcase the company's commitment to sustainability (Lueg et al., 2016).

2.3 Central and Eastern Europe: economic challenges and compliance issues

The literature identifies significant disparities in ESG disclosures between 'old' and 'new' EU Member States (those that joined the EU before or after 2004), with CEE countries generally showing lower quality reporting (Janicka and Sajnóg, 2023; Tozser et al., 2024). Bochenek, 2020 highlights a strong correlation between GDP per capita and corporate reporting, suggesting that wealthier countries produce more comprehensive reports, suggesting that the disparity in reporting quality could be linked to different levels of economic development, with the emerging stock exchanges and evolving economies of CEE countries compared to 'old' EU Member States limit comprehensive reporting. Moreover, the transposition of the NFRD varied significantly across CEE countries, leading to disparities in reporting due to different implementation strategies, enforcement levels, and unique national contexts (Bochenek, 2020; Dragija Kostic et al., 2017; Dumitru et al., 2019; Janicka and Sajnóg, 2023; Schröder, 2022). Countries with well-established governance systems are better at enforcing compliance and encouraging comprehensive reporting (Bochenek, 2020), while CEE countries showed limited preparation for the directive (Dumitru et al., 2019). Bulgaria and Romania, for instance, transitioned from centrally governed socialist regimes to market economies upon entering the EU in 2007, but this did not lead to stable development, partly due to modest financial markets and poor social and political stability (Aureli et al., 2020). However, while CEE countries lag behind "old" EU Member States in non-financial reporting, NFRD still proved effective in this environment as evidenced by better disclosures in Balkan EU members compared to non-EU counterparts (Tocev et al., 2022).

The research found that the NFRD did not immediately improve ESG performance, with significant progress occurring only in the second and third years after its implementation, suggesting a period of adaptation was necessary for companies (Aluchna et al., 2022; Belenesi et al., 2021; Dragomir et al., 2024; Dumitru et al., 2019; Hategan et al., 2021). Poland produces the largest number of non-financial reports among CEE countries (Krasodomska et al., 2020), with this number significantly increasing as a result of the Directive's implementation (Matuszak and Rózanska, 2021), followed by Hungary (Arraiano and Hategan, 2019). However, the quality is found to remain pretty low, with the overall disclosure quality increasing from a low to a medium level after the implementation of the directive (Borowicz and Czerepko, 2023; Hategan et al., 2021; Lippai-Makra et al., 2022), with a predominance of positive impacts disclosed, and a tendency to report more on risk identification than on actions taken to mitigate these risks (Dragomir et al., 2024; Tamm and Gurvits-Suits, 2023), and a lack in qualitative aspects (Primec and Belak, 2022).

Major obstacles to non-financial reporting development in transition countries transition countries include misunderstandings of the concept, lack of qualified staff, difficulty in perceiving direct business benefits, poor government incentives, insufficient time, and limited financial resources (Krasodomska et al., 2020; Primec and Belak, 2022). Executives are hesitant to prioritize sustainability due to concerns about high costs, leading to selective disclosure of sustainability efforts (Jovanovič and Jovanović, 2022; Klimczak et al., 2023). However, non-financial reporting awareness, although generally low among accounting specialists, improves significantly with ESG training, leading to more positive attitudes and a better understanding of mandatory and environmental reporting (Janicka and Sajnóg, 2023; Krasodomska et al., 2020).

A significant challenge is the predominance of smaller firms in the region, which results in a limited number of companies meeting the directive's criteria and thus hinders widespread corporate compliance and the fostering of a broad culture of sustainability reporting. Indeed only 50 firms are concerned in Slovenia according to (Primec and Belak, 2022), and 200 in Croatia according to (Černe et al., 2017), given the small scale of businesses, even thought it had broaden the definition of PIEs to include credit institutions, and (re)insurance companies but also financial institutions like leasing companies, investment funds, and entities of strategic interest to Croatia, as determined by the government. Hungary has further narrowed the scope with additional thresholds on net sales revenue and total assets, reducing the number of entities subject to non-financial reporting and potentially compromising the directive's effectiveness in achieving transparency (Lippai-Makra et al., 2022).

2.4 Discussion

In synthesizing the literature on the national implementations on the NFRD, key themes and categories (first-order concepts) are identified through a structured approach, analogous to axial coding, and will now employ abductive reasoning to iteratively refine our understanding of the relationships and patterns within the data. This process allows us to examine the causal contributions of various conditions and understand the factors influencing non-financial reporting across European countries. One of the significant factors identified is national culture, which plays a crucial role in shaping corporate behaviors and reporting practices. To explore this further, comparative analysis that accounts for the potential impact of cultural dimensions on ESG reporting across different European nations is conducted.

Cultural factors have been analyzed from various perspectives, but the scientific debate is largely lead by the widespread adoption of Hofstede's theoretical framework (Hofstede, 1980). Hofstede, 1980 define culture as 'the collective programming of the mind which distinguishes the members of one human group from another.' Hofstede identified the following cultural dimensions: power distance (PD), uncertainty avoidance (UA), individualism versus collectivism (IDV), masculinity versus femininity (MAS), and long- versus short-term orientation (LTO), and indulgence versus restraint (IVR) (added at a later stage of his research (Hofstede et al., 2014). Academics have used the model to assess the relationship between cultural factors and reporting harmonization and standardization (Hąbek, 2018; Mittelbach-Hörmanseder et al., 2021; Panfilo and Krasodomska, 2022; Pizzi et al., 2022; Velte, 2023a; Zanellato and Tiron-Tudor, 2022). Research suggests that no single cultural profiles is solely responsible for high disclosure levels. Instead, it is the combination of different cultural dimensions that leads to varying levels of corporate disclosure (Zanellato and Tiron-Tudor, 2022).

[Insert Figure 1.2 about here]

From the literature it appears that the main driver for high quality reporting is the pre-existing reporting culture and strong implementation of the directive (e.g., France, Spain, Scandinavian countries). A strong implementation alone however does not seem to lead to high quality reporting if the pre-existing reporting culture was low (e.g. Italy).

Including Hofstede's cultural dimensions, Scandinavian countries exhibit cultural traits of low power distance, high individualism, and low masculinity, creating a conducive environment for transparent and ethical business practices, fostering egalitarianism, self-reliance, and a preference for cooperation and quality of life over competition. Low PD indicates a society with more equal power distribution and less pronounced hierarchies, fostering participatory decision-making, inclusiveness, transparency, and accountability, which are essential for embracing and integrating ESG initiatives and ensuring comprehensive ESG reporting and practices that uphold stakeholder accountability. High IDV indicates a society that values individual rights, personal achievements, and self-expression, fostering a sense of personal responsibility and ethical behavior that supports ESG practices aligned with personal values, while companies prioritize stakeholder engagement, driving better

ESG practices to meet the expectations of employees, customers, and the community. Low MAS indicates a society that values quality of life, nurturing, and cooperation over competition and achievement, prioritizing well-being and ESG-aligned practices, fostering a collaborative environment that enhances ESG implementation and reporting, and balancing financial performance with social and environmental goals for comprehensive ESG integration. These cultural traits foster environments conducive to robust ESG reporting, aligning with the observed high-quality disclosures.

Before the directive, Germany, while not having a strong legal framework before the directive, did already have high levels of ESG reporting. This is attributed to its cultural traits similar to Scandinavian countries of low PD and high IDV, which foster a reporting culture. In this cultural context, normative mechanisms, which rely on cultural and societal norms, are more effective than coercive mechanisms, which depend on legal enforcement.

Economic development and governance systems also play crucial roles in the effectiveness of non-financial reporting. The disparities between 'old' and 'new' EU Member States, particularly in CEE countries, highlight the impact of economic conditions on reporting practices. These countries also present specific cultural dimensions of high UAI and low IVR. High UAI cultures, resistant to change, may struggle to adopt new reporting practices or innovative ESG approaches, resulting in lower quality and comprehensiveness. Additionally, low IVR cultures, focusing on financial stability and traditional business practices, may prioritize short-term financial performance over long-term sustainability initiatives due to perceived lack of immediate benefit. These cultural factors contribute to the lower quality of reporting observed in CEE countries.

3 Firm-level cultural dynamics: Balancing representation and power in Corporate Governance

Having explored the impact of country-specific cultural and legal frameworks on the implementation of Directive 2014/95/EU, this section shifts focus to how firm-specific cultural values within corporate governance structures influence the effectiveness of sustainability disclosures required by the directive. Corporate governance structures within firms—such as board size, independence, gender diversity, and the presence of CSR committees— significantly impact ESG reporting practices (Hamed et al., 2022; Radu et al., 2023; Velte, 2023a). These internal factors, driven by the unique cultural values and priorities of its executive leadership, significantly impact how companies approach transparency and accountability in their sustainability reports (García-Sánchez et al., 2023). This section examines how the interplay of these governance elements and managerial attitudes drives ESG reporting, complementing the broader regulatory and cultural landscape.

3.1 Gender diversity on boards: varying impacts across cultural contexts

According to prior literature reviews, research finds that gender diversity on the board of directors enhances the quality of ESG disclosures (Radu et al., 2023; Velte, 2023a). Women, who often prioritize quality of life over material success (Hofstede et al., 2014), are more attentive to environmental responsibilities and societal well-being. Research in international and EU-wide settings argue that based on stakeholder and resource dependence theories, their presence contributes to bringing diverse skills, perspectives, leadership styles, and external connections, which enhances decision-making and increases accountability and transparency in ESG matters due to their greater concern for societal welfare and sustainability (Bigelli et al., 2023; Chimonaki et al., 2024; García-Sánchez et al., 2023; Nicolò et al., 2022; Pizzi et al., 2023; Velte, 2023a, 2023b). However, country-specific analyses yield different results.

Baalouch et al., 2019 confirm that in the French context, characterized by medium-low level of masculinity (Hofstede et al., 2014) and a regulatory environment requiring gender parity on boards, the positive significant relationship is present. Similarly in Portugal, with a low level 31 of MAS, Lemos et al., 2021 find that the proportion of women directors on the board has a significant positive influence on non-financial mandatory disclosure, with companies with gender-diverse boards show higher compliance with non-financial reporting requirements than those with male-dominated boards. The results are however less clear in the Italian and German contexts, characterized by higher levels of masculinity (Hofstede et al., 2014). In the Italian context, Caputo et al., 2020 found no significant relationship between the two, and Beretta et al., 2023 argue that while a high percentage of women might not improve ESG disclosure quality, an increase in female representation from a low to a medium level is associated with better ESG reporting. Under the German two-tier system, Gerwing et al., 2022 found no significant relationship between the proportion of women on the executive board and the quality of sustainability reporting; however, they found a positive significant one with the supervisory board, responsible for appointing the executive board and auditing the content of the nonfinancial statement under § 171 Stock Corporation Act.

In the context of the research, those findings suggest that the positive effect of gender diversity on the board on ESG disclosure is embedded in the cultural context of low masculinity in the country.

3.2 Board size and independence: more voices, better outcomes?

Similarly, as stakeholder theory posits that organizations should manage their business with an emphasis on the interests of all stakeholders; and as a larger board encompasses a wider range of stakeholders, facilitating more proactive stakeholder management and resulting in enhanced ESG reporting, board size is found to have a positive relationship with ESG disclosures (Dias et al., 2017; García-Sánchez et al., 2023; Mio et al., 2020;

Radu et al., 2023; Rep et al., 2022). Mio et al., 2020 even find that this effect is further reinforced after the implementation of the NFRD, suggesting that board size plays an even bigger role in firms' choice to disclose more ESG information in a regulated environment. However, Nicolò et al., 2022 note that beyond a certain threshold, the benefits of larger boards may be outweighed by communication and coordination problems, potentially hindering the quality and thoroughness of ESG disclosures. This supported by Beretta et al., 2023, which finds that while larger boards increase the length and content of non-financial disclosures, they also might require more consensus between the diverse viewpoints within the board leading to more strategic, selective disclosures, which can limit the completeness of the information presented. This observation suggests a need for an optimal board size that balances diversity and manageability to facilitate effective governance.

Larger boards often include more independent directors, bringing together non-executive and independent members who provide diverse perspectives and enhance oversight. These directors, not involved in daily management and without significant ties to the company's ownership, can improve governance practices and ensure comprehensive ESG disclosures. Their independence from internal biases allows them to monitor and influence management effectively, addressing the interests of all stakeholders (García-Sánchez et al., 2023; Nicolò et al., 2022; Pizzi et al., 2023; Radu et al., 2023).

However, it is important to consider regional differences. For instance, at the French level, the effect is inverted. Baalouch et al., 2019 suggest that the reluctance of external independent directors to promote non-financial disclosure may stem from their limited understanding of the reliability of such information. They might also fear that disseminating inaccurate information could heighten the risks to their reputation, as they are not directly involved in the day-to-day operations and cannot fully verify the disclosed data.

Similarly, Dias et al., 2017 find that in the context of Portugal, the effectiveness of independent directors is limited by the high ownership concentration and prevalent CEO duality typical of their country, reducing their ability to positively influence ESG disclosures. Portuguese companies typically exhibit high ownership concentration, where large shareholders, often families, have significant control over the company.

In the context of the research on the impact of legal framework and cultural factors, those findings suggest that board independence has a negative impact on ESG disclosure in cultural context with high PD (France and Portugal), indicating a hierarchical society where power is centralized and authority is not questioned. Independent directors may find it challenging to exert their influence in such an environment because their role is often overshadowed by the entrenched power of the dominant shareholders. As a result, their ability to drive comprehensive ESG disclosures is diminished. The effect is further reinforced in a context of low individualism (Portugal), leaning towards a collectivist society where group cohesion and loyalty are paramount. In collectivist cultures, decision-making processes tend to emphasize consensus and maintaining harmony. Independent directors, who are expected to provide objective oversight, may be reluctant to push for changes that disrupt established practices or challenge the majority shareholders. The collectivist nature may result in a tendency to align with the interests of the dominant shareholders, which can lead to selective ESG disclosures that favor the controlling families' perspectives and interests over broader stakeholder transparency.

3.3 Power concentration in Governance: challenges to transparency and the role of regulation

The presence of CEO duality, where the CEO also serves as the board chairman, is generally one of the main drivers negatively affecting the transparency of non-financial reporting. CEO duality refers to a governance structure where the roles of the CEO and board chairman are held by the same individual. Prior studies indicate that CEO duality limits effective accountability regarding unfavorable information due to the CEO's vested interest in the firm's performance. This governance structure tends to limit corporate interest towards sustainability and reduces the representation of minority stakeholders, thereby generally having a negative impact on the transparency of non-financial reporting (Caputo et al., 2020; García-Sánchez et al., 2023). Bigelli et al., 2023 found that before the implementation of the EU Non-Financial Reporting Directive, CEO duality negatively affected ESG scores. However, the directive, which mandated a focus on ESG issues, mitigated this negative impact, making the influence of CEO duality on ESG disclosures insignificant. This suggests that mandatory reporting can counteract the adverse effects of CEO duality on ESG disclosures. The directive led to higher ESG scores across firms, reducing the discrepancy caused by CEO duality (Bigelli et al., 2023).

Ownership structure, whether the shares of the company are concentrated or dispersed among shareholders, plays a crucial role in the extent of disclosure of additional information. Studies indicate that closely held firms, where ownership is concentrated among a few insiders, tend to perform lower on the disclosure level of ESG. This is attributed to possible agency problems, where the interests of the closely held shareholders may not align with the broader stakeholder interests, resulting in lower ESG disclosures (Rezaee et al., 2023). Interestingly, Gerwing et al., 2022 find that while closely held shares tend to negatively impact the quality of disclosure due to reduced external pressures, the implementation of robust corporate governance mechanisms can enhance the overall quality of mandatory sustainability reporting. This interplay between ownership structure and corporate governance highlights the importance of comprehensive governance frameworks in promoting transparency and accountability in corporate reporting. Effective governance mechanisms (e.g. sustainable remuneration, ESG committee, engagement in ESG initiatives, external assurance) can mitigate the negative impacts of concentrated ownership by ensuring that broader stakeholder interests are considered in corporate decision-making.

The negative impact of CEO Duality and high concentration ownership on ESG Disclosures is mitigated by the legal requirement and governance mechanisms, further emphasizing the importance of the upcoming CSRD that will make external assurance mandatory.

4 Preparing for enhanced sustainability reporting: key themes under the CSRD

The EU Commission also published in 2017 'guidelines on non-financial reporting' (European Commission, 2017) to support companies in their implementation of the requirements. While non-binding they provide clarification with regards to key principles for ensuring the quality and effectiveness of the NFRD. After a public consultation period from February to June 2020 to gather stakeholder input (European Commission, 2020), the European Parliament adopted in 2022 the Corporate Sustainability Reporting Directive (CSRD) (European Commission, 2022). The new directive will significantly increase the scope, from 11,000 to around 50,000 expected to be concerned by the new reporting requirements by 2024, as well as stricter reporting standards developed by the European Financial Reporting Advisory Group (EFRAG).

As such the following section identifies another strain of themes starting to emerge relating to the specific content of the reports, as CSRD will reinforce that overlooked aspect in the initial NFRD.

4.1 Addressing materiality: from NFRD ambiguities to CSRD clarity

A critical aspect of the Directive is the concept of materiality. In corporate reporting, it consists in ensuring that disclosures focus on the most significant issues for the organization, society, and the environment, making the reports both comprehensive and concise, and addressing stakeholder needs through regular assessments of relevant factors (European Commission, 2017).

Despite its importance, the NFRD lacked a clear definition or guidance on materiality and its assessment, leading to inconsistencies in how Member States implement and interpret these requirements (Aureli et al., 2018; García-Sánchez et al., 2023; La Torre et al., 2020). A perfect example is raised by Raith, 2023, where Austria and Germany have adopted different interpretations of the NFRD's materiality provision based on language nuances. Austria's NaDiVeG uses an 'additive' approach, requiring disclosure of societal and environmental impacts, while Germany's CSR-RL-UG takes a 'cumulative' approach, mandating disclosure only if these impacts also affect the company's performance. Consequently, Gerwing et al., 2022 found significant heterogeneity in German firms' sustainability reporting, with many not fully addressing material aspects, leading to varied and often low-quality disclosures. In the Spanish context, Ruiz et al., 2021 found in their study of SOEs that while the quality of materiality disclosure is generally low, larger entities (> 250 employees) and companies that use external assurance have significantly higher materiality disclosure levels. This is in line with Mazzotta et al., 2020 and Lopez et al., 2023, both finding significantly high commitment to materiality assessment in Italy and Spain, countries where external assurance is mandatory.

The introduction of the CSRD aims to address these issues, by making external assurance mandatory and clearly defining the concept of double materiality in the European Sustainability Reporting System (Baumuller and Sopp, 2022; Hummel and Jobst, 2024; Mezzanotte, 2023) as the integration of both financial materiality (outside-in), which focuses on external factors affecting enterprise value for investors, and impact materiality (inside-out), which considers the entity's societal and environmental impact (Fiandrino et al., 2022; La Torre et al., 2020; Lombardi, Cosentino, et al., 2022). As such newer studies are analysing the impact of the new definition of double materiality.

Initial research of indicators of double materiality found that in the first years following the introduction of the concept (2019-2021), only a small percentage of companies referenced double materiality or its pillars in their reports (De Cristofaro and Gulluscio, 2023). Moreover those that did include it often lacked visual aids or references to the EU's regulatory framework. Western European countries like Germany, Italy and Spain were found to be more proactive in adopting double materiality (De Cristofaro and Gulluscio, 2023). In a qualitative document analysis, Miettinen, 2024 compared the situation of companies in Germany (Bayer), Denmark (Novo Nordisk), France (Sanofi) and Belgium (UCB). The analysis revealed that Nordisk and Sanofi had conducted a double materiality assessment for the first time in 2022 while UCB and Bayer have not yet carried it out.

Although little research is available on the subject in the CEE area, the studies that do exist present coinciding results. In Poland,despite being relatively experienced in non-financial reporting compared to its neighbors, there are still significant gaps in meeting the new ESRS requirements (Broniewicz et al., 2024). Similarly, in Romania, Dragomir et al., 2024 found that in their 2022 reports, almost all companies in the sample disclosed information about their materiality assessments, demonstrating a strong recognition of its importance. However, only five companies (25% of the sample) explicitly mentioned double materiality.

Collectively, these findings suggest that while the concept of double materiality is gaining recognition across Europe, it remains an emerging practice with significant variations in its application. There is a notable lack of empirical knowledge on how companies practically implement double materiality, including the methods and processes they use for assessment and reporting (Baumuller and Sopp, 2022; De Cristofaro and Gulluscio, 2023; Dragomir et al., 2024). Further research is needed to explore the application and disclosure of the materiality principle, especially in understanding how companies interpret and apply it in sustainability reporting. This is crucial, as the strategic selection of sustainability topics by companies can significantly influence stakeholder interests, with the potential for companies' interpretations of materiality to undermine these interests (Baumuller and Sopp, 2022; La Torre et al., 2020; Ruiz et al., 2021). Additionally, there is a clear need to better conceptualize and operationalize materiality, particularly ecological and social materiality (Baumuller and Sopp, 2022; Miettinen, 2024), as the current understanding and practices surrounding materiality are still evolving, and further research is essential to establish a more robust and concrete framework.

4.2 Strategic and forward-looking reporting under the CSRD

Another critical aspect is the fact that the reports should be strategic and forward-looking. As such they should provide insights into a company´s business model, strategy and its implementation, explain the short-term, medium-term and long-term implications of the information reported, proposing a fair and balanced view of both favorable and unfavorable aspects, with targets and benchmarks presented both in qualitative and quantitative terms (European Commission, 2017).

The guidelines encourage companies to provide comprehensive information about their business models, strategies, and the implementation of these strategies as part of the forward-looking information (Korca, 2022; Simoni et al., 2022). Companies are expected to include applicable time frames (retrospective and forward-looking information is required) (Baumuller and Grbenic, 2021), elucidating the short-term, medium-term, and long-term implications of the reported information (Primec and Belak, 2022), thereby providing a comprehensive understanding of their trajectory and operational context (Carungu et al., 2021) and on how their resources have evolved (Scandurra and Thomas, 2023).

Incorporating strategies aligned with the sustainability into business models necessitates a long-term perspective due to the lack of an immediate link to utilitarian goals (Pizzi et al., 2022). Combining business model information with financial outcomes and sustainability information using forward-looking targets helps investors evaluate future earnings, with forward-looking data serving as a leading indicator, showcasing long-term potential (Glaveli et al., 2023; Simoni et al., 2022). However, research shows a lack of consensus on the definition, components, and function of the business model, indicating it is still under social construction in this context (Bini et al., 2023; Gerwing et al., 2022). Moreover, firms often avoid the disclosure of sensitive forward-looking information about their business model due to business secrecy risks, which leads companies to limit their disclosures to general information to avoid litigation risks (Bini et al., 2023; Nicolo et al., 2021).

Research and professional bodies both recognize a growing trend towards time-oriented, forward-looking reporting (Dencic-Mihajlov and Spasic, 2016; Heichl and Hirsch, 2023). Stakeholders have pointed out during the consultation process of the NFRD that current financial reporting's historical focus inadequately captures future value creation, making forward-looking sustainability data essential for assessing a company's future viability and environmental impacts (Fiandrino et al., 2022). Additionally, the new CSRD proposal, suggests that companies should include both qualitative and quantitative forward-looking information (Fiandrino et al., 2022; Primec and Belak, 2022), offering a clearer picture of their objectives and operational targets. As such research on forward-looking information include quality variables, such as tone and perspective (Agostini et al., 2022). A focused examination of Italian companies found a clear positive relationship between ESG coverage and the length of non-financial disclosures with narrative features, such as forward-looking orientation (Beretta et al., 2023). Moreover, readability improves when sustainability-related information is forward-looking, aligning with the incremental information approach, as it aids investors in understanding firm performance and future earnings (Beretta et al., 2023).

Based on the findings from Heichl and Hirsch, 2023, while there is a general trend towards future-oriented reporting, there are notable differences in forward-looking disclosures among France, Germany, Italy, and Sweden. Germany and Sweden, both characterized by LTO cultures, are the leaders in forward-looking reporting, with 28.84% and 31.23% of their disclosures focused on future-oriented topics, respectively. In contrast, France and Italy report significantly less forward-looking information, with only 21.90% and 21.73% of their disclosures oriented towards the future.

Regarding the integration of sustainability into company strategies, (Pizzi et al., 2022) explored how cultural dimensions influence SDG reporting practices. Their study reveals that cultural factors like LTO and IVR positively impact SDG reporting, reflecting the forward-looking nature of the 2030 Agenda, while IDV and MAS negatively affect it. Notably, Scandinavian firms, recognized for their socially responsible behavior, lead in SDG reporting. However, their analysis reveals that while companies may report on SDGs, this reporting does not necessarily indicate a genuine integration of these goals into their core business strategies. The research underscores that SDG reporting is often influenced by cultural dimensions and personal commitments rather than by a deep-rooted integration into business models. This suggests that the inclusion of SDGs in non-financial reports is frequently more about voluntary disclosure driven by external factors, rather than a reflection of intrinsic business practices aligned with sustainable development goals.

Despite a trend towards more future-oriented disclosures, this area remains a relatively understudied and requires further exploration. Future research should focus on understanding the impact of these disclosures on stakeholder decision-making, examining the divergence in perceptions between the users and preparers around the business model concept and its content (Bini et al., 2023), identifying best practices for balancing transparency with the risks of disclosing sensitive information (Cuomo et al., 2024), and exploring how cultural dimensions influence the genuine adoption of sustainable practices in business models (Pizzi et al., 2022; Simoni et al., 2022). As the CSRD evolves, it is essential to monitor how companies adapt their reporting strategies to meet these requirements, ensuring the intended goals of enhanced transparency and accountability are fully realized (Aboud et al., 2024).

4.3 Additional research directions: long-term analysis and scope expansion

The transition from the NFRD to the CSRD presents significant opportunities for future research, particularly in understanding how the enhanced regulatory framework will influence corporate sustainability practices across Europe. Several areas require further investigation to support this transition and contribute to the evolving body of knowledge on corporate sustainability reporting.

While the short-term effects of non-financial reporting directives are well-studied, there is a notable lack of longitudinal research on their long-term impact. Most studies have focused on initial compliance, leaving gaps in understanding how these practices evolve and influence transparency and sustainability over time (Fiechter et al., 2022; Hummel and Rotzel, 2019). Several studies emphasize the need for longitudinal research to study the evolutionary process of corporate responses to sustainability reporting regulation, revealing trends that short-term studies may overlook (Aluchna et al., 2022; Esteban-Arrea and Garcia-Torea, 2022; Mazzotta et al., 2020; Raucci and Tarquinio, 2020). As time progresses and more data becomes available, it will be possible to conduct longitudinal research, allowing for a deeper analysis of the long-term effects and trends related to corporate behavior, transparency, and sustainability practices.

While there have already been calls to include smaller entities alongside the large ones that are predominantly studied (Arraiano and Hategan, 2019; Hategan et al., 2021; Kühn et al., 2014; Mion and Adaui, 2019), notably to investigate the trickle-down effect of regulation on small and medium-sized firms (García-Sánchez et al., 2023), this has become even more pertinent with the expansion of the scope of CSRD. The directive significantly increases the number of companies required to comply, now including large unlisted and SME listed companies. Future research should leverage the increased sustainability data from the CSRD to gain a deeper understanding of the role of these smaller entities in the transition to sustainability (Hummel and Jobst, 2024), but present the challenge of aggregating data from a larger and more diverse set of companies into meaningful sustainability measures.

5 Conclusion

The objective of this literature survey is to synthesize and critically examine the potential cause-and-effect relationships emerging from the body of literature on the implementation of the EU NFRD across various European countries. A central theme identified in this survey is the non-prescriptive nature of the NFRD, which has resulted in diverse interpretations and implementations across the EU, subsequently affecting the overall quality and comparability of non-financial reports. This theme guided the selection of papers reviewed, which were analyzed to document the varied configurations of conditions associated with the observed outcomes.

In the second section, the review reveals the national implementation variations, highlighting contradictory findings and the evolving nature of the regulations within the EU. The effectiveness of the NFRD is significantly influenced by pre-existing reporting cultures, national implementation strategies, and cultural dimensions. In countries with robust pre-existing practices and supportive cultural traits, higher quality reporting is achieved. The findings suggest that normative mechanisms generally prove more effective than coercive ones. For regions with lower initial reporting quality, such as CEE countries, targeted support and phased implementation are recommended to improve compliance and reporting standards. This underscores the importance of considering cultural and economic contexts in developing effective non-financial reporting frameworks across the EU's diverse legal landscapes.

The third section identifies the differential impact of corporate governance practices, depending on the cultural context in which companies operate. For instance, while gender diversity on boards positively influences ESG disclosure at the EU level, this effect is confirmed only in countries with lower masculinity indices in country-specific studies. Similarly, the positive influence of board independence observed at the EU level is reversed in countries with high levels of power distance and collectivism. Furthermore, the negative impact of CEO duality and high concentration of ownership on ESG disclosures is mitigated by legal requirements and governance mechanisms, emphasizing the significance of the forthcoming CSRD in mandating external assurance.

In the fourth section, the review delves into the implications of the upcoming CSRD, identifying key themes as the EU transitions from the NFRD. The CSRD represents a substantial evolution in corporate reporting, introducing stricter requirements, broader scope, and mandatory external assurance, which are expected to enhance the transparency, comparability, and overall quality of non-financial reports across Europe. This regulatory shift also highlights the increasing importance of materiality, forward-looking reporting, and the integration of sustainability into corporate strategies, reflecting a more comprehensive approach to corporate accountability. However, several challenges persist in this evolving regulatory landscape. For example, while forward-looking reporting under the CSRD is crucial for understanding a company's strategy and long-term viability, its implementation faces challenges, including inconsistencies in defining business models, concerns over business secrecy, and cultural factors that influence the integration of sustainability goals.

In conclusion, this review underscores the complex interplay between regulatory frameworks, cultural contexts, and corporate governance in shaping ESG reporting practices across Europe. The transition from the NFRD to the CSRD marks a critical juncture in the EU's efforts to lead global sustainability standards, offering both challenges and opportunities for enhancing corporate transparency and accountability. As the CSRD is implemented, ongoing research will be essential to monitor its impact, identify best practices, and ensure the directive achieves its intended goals of fostering a more sustainable and transparent business environment across Europe. Future research should emphasize longitudinal analyses to understand the long-term effects of these regulations, explore the implications of the CSRD on smaller firms, and investigate the evolving role of cultural dimensions in shaping corporate sustainability practices. By doing so, scholars and practitioners can deepen the understanding of how legal, societal, and cultural factors influence the effectiveness of sustainability reporting, thereby supporting the EU's broader objectives of promoting sustainability and corporate responsibility on a global scale.

6 Figures and Tables

Co-occurrence of countries being studied together

Country	Total Obs.	Single Obs.	Slovenia	Croatia	Hungary	Slovakia	Lithuania	Estonia	Latvia	Romania	Bulgaria	Poland	Sweden	Belgium	Netherlands	Finland	UK	Denmark	Austria	Ireland	Italy	France	Germany	Spain
Spain	11	7												1	2	1				1	4	1	2	
Germany	8	2											1	1	2	1			1	1	5	2		2
France	8	3								1	1	1	2	2	2	1	3	1		1	3		2	1
Italy	29	20											1	1	2	1	2			1		3	5	4
Ireland	1	0												1	1	1					1	1	1	1
Austria	1	0																					1	
Denmark	4	1										1	3		1	2	1					1		
UK	8	4								1	1	1	1	1	1			1			2	3		
Finland	3	0											2	1	1			2		1	1	1	1	1
Netherlands	3	0										1	1	1		1	1	1		1	2	2	2	2
Belgium	2	0								1	1				1	1	1			1	1	2	1	1
Sweden	5	1										1			1	2	1	3			1	2	1	
Poland	8	6	1	1	1	1	1	1	1	1	1		1		1		1	1				1		
Bulgaria	2	0	1	1	1	1	1	1	1	2		1		1			1					1		
Romania	5	3	1	1	1	1	1	1	1		2	1		1			1					1		
Latvia	1	0	1	1	1	1	1	1		1	1	1												
Estonia	2	1	1	1	1	1	1		1	1	1	1												
Lithuania	1	0	1	1	1	1		1	1	1	1	1												
Slovakia	1	0	1	1	1		1	1	1	1	1	1												
Hungary	4	2	2	2		1	1	1	1	1	1	1												
Croatia	5	2	3		2	1	1	1	1	1	1	1												
Slovenia	4	1		3	2	1	1	1	1	1	1	1												
Portugal	5	5																						
Greece	1	1																						

Figure 1.1: Frequency of European countries as research interest in the literature
This figure illustrates the frequency with which specific European countries are examined in the selected literature. The figure does not include studies that analyzed all European countries collectively but focuses on those that explore specific groupings. The table on the left summarizes the total number of studies that focus on each country, whether alone or with another country (Total Obs.) and the number of studies where each country is analyzed independently (Single Obs.). The heatmap on the right visualizes the co-occurrence of country pairs in these studies, with darker shades indicating a higher frequency of joint analysis.

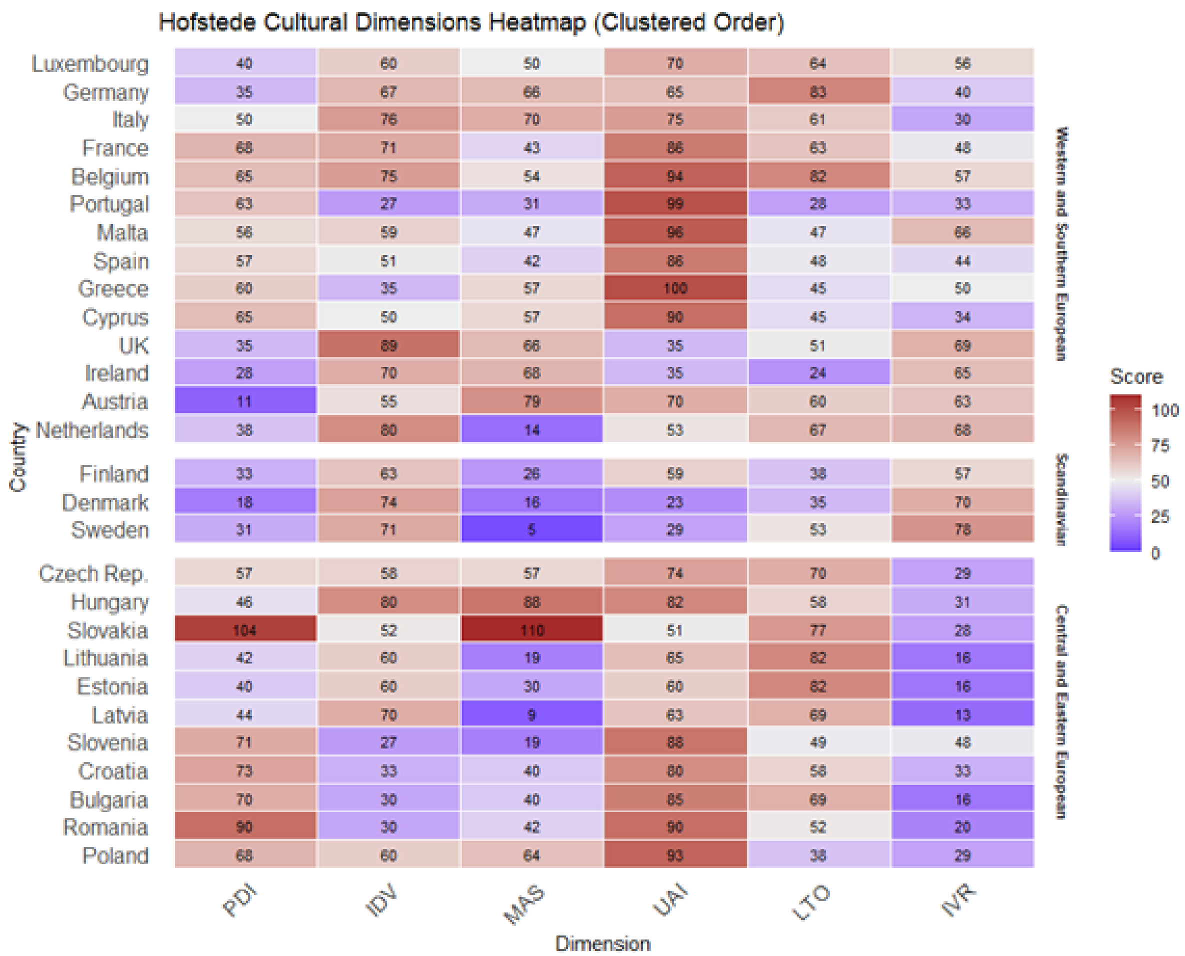
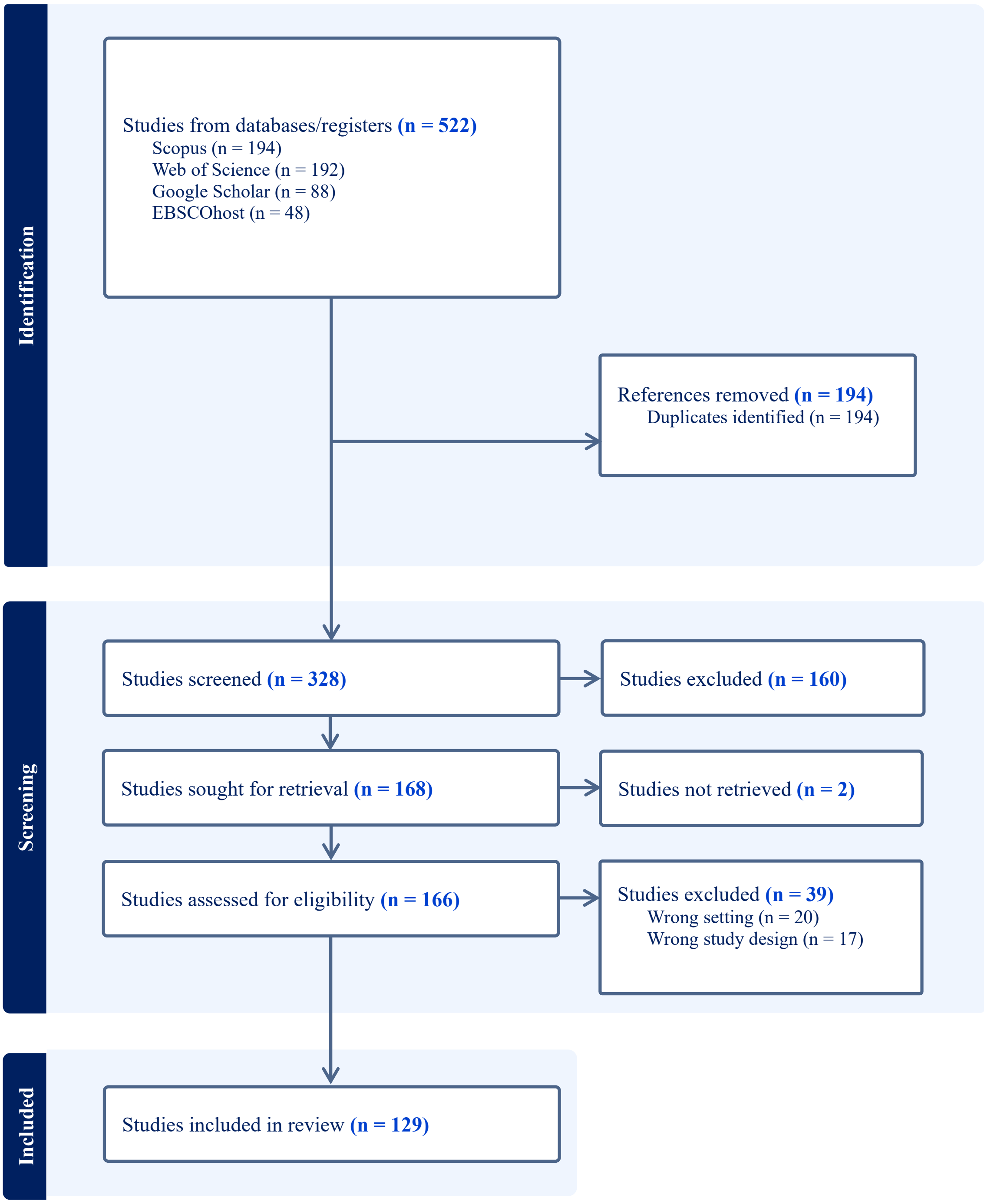


Figure 1.2: Hofstede cultural dimensions heatmap (Clustered Order)
This heatmap illustrates Hofstede country index values for six culture dimensions for European countries, organized by region: Western and Southern European, Scandinavian, and Central and Eastern European. The countries within each region are ordered according to hierarchical clustering based on their cultural dimensions, enhancing the visualization of regional cultural patterns.

7 Appendices

Appendix A - PRISMA



8 References

Aboud, A., Saleh, A., & Eliwa, Y. (2024). Does mandating ESG reporting reduce ESG decoupling? Evidence from the European Union's Directive 2014/95 [number-of-pages: 16 publisher-place: Hoboken publisher: Wiley Web of Science ID: WOS:001082683300001]. *BUSINESS STRATEGY AND THE ENVIRONMENT*, *33*(2), 1305–1320. https://doi.org/10.1002/bse.3543 (cit. on p. 37).

Agostini, M., Costa, E., & Korca, B. (2022). Non-Financial Disclosure and Corporate Financial Performance Under Directive 2014/95/EU: Evidence from Italian Listed Companies. *ACCOUNTING IN EUROPE*, *19*(1), 78–109. https://doi.org/10.1080/17449480.2021.1979610 (cit. on pp. 12, 17, 36).

Aguado-Correa, F., de la Vega-Jimenez, J. J., Lopez-Jimenez, J. M., Padilla-Garrido, N., & Rabadan-Martin, I. (2023). Evaluation of Non-financial Information and Its Contribution to Advancing the Sustainable Development Goals within the Spanish Banking Sector: European Research on Management and Business Economics [[Online; accessed 2024-05-02]]. *European Research on Management and Business Economics*, *29*(1). https://doi.org/10.1016/j.iedeen.2022.100211 (cit. on p. 18).

Aluchna, M., Roszkowska-Menkes, M., & Kaminski, B. (2022). From talk to action: The effects of the non-financial reporting directive on ESG performance [number-of-pages: 25 publisher-place: Bingley publisher: Emerald Group Publishing Ltd Web of Science ID: WOS:000877186300001]. *MEDITARI ACCOUNTANCY RESEARCH*, *31*(7), 1–25. https://doi.org/10.1108/MEDAR-12-2021-1530 (cit. on pp. 12, 23, 38).

Andersson, F. N. G., & Arvidsson, S. (2023). Understanding, Mapping and Reporting of Climate-Related Risks among Listed Firms in Sweden: Climate Policy. *Climate Policy*, *23*(8), 945–958. https://doi.org/10.1080/14693062.2022.2116383 (cit. on p. 22).

Anguiano-Santos, C., & Rodriguez-Entrena, M. (2024). Sustainability Reporting in Focus: Analysing Spanish Transposition of the Non-financial Reporting European

Directive in the Agri-food Sector: Agricultural and Food Economics. *Agricultural and Food Economics*, *12*(1). https://doi.org/10.1186/s40100-024-00303-z (cit. on p. 18).

Araya, H., & Sierra-García, L. (2021). Assurance on non-financial information in Spain. *CONTADURIA UNIVERSIDAD DE ANTIOQUIA*, *79*, 13–37. https://doi.org/10.17533/udea.rc.n79a01 (cit. on p. 18).

Arif, M., Gan, C., & Nadeem, M. (2022). Regulating non-financial reporting: Evidence from European firms' environmental, social and governance disclosures and earnings risk. *MEDITARI ACCOUNTANCY RESEARCH*, *30*(3), 495–523. https://doi.org/10.1108/MEDAR-11-2020-1086 (cit. on p. 12).

Arraiano, I. G., & Hategan, C. D. (2019). The Stage of Corporate Social Responsibility in EU-CEE Countries [number-of-pages: 14 publisher-place: Rome publisher: European Center Sustainable Development Web of Science ID: WOS:000488623700034]. *EUROPEAN JOURNAL OF SUSTAINABLE DEVELOPMENT*, *8*(3), 340–353. https://doi.org/10.14207/ejsd.2019.v8n3p340 (cit. on pp. 23, 38).

Aureli, S., Magnaghi, E., & Salvatori, F. (2018). The transposition of the non-financial reporting directive in the UK, France and Italy [publisher: University of Leicester] (cit. on pp. 13, 17–20, 32).

Aureli, S., Salvatori, F., & Magnaghi, E. (2020). A Country-Comparative Analysis of the Transposition of the EU Non-financial Directive: An Institutional Approach: Accounting, Economics, and Law: A Convivium. *Accounting, Economics, and Law: A Convivium*, *10*(2), 1–30 (cit. on pp. 12, 13, 18, 19, 23).

Baalouch, F., Ayadi, S., & Hussainey, K. (2019). A study of the determinants of environmental disclosure quality: Evidence from French listed companies. *JOURNAL OF MANAGEMENT AND GOVERNANCE*, *23*(4), 939–971. https://doi.org/10.1007/s10997-019-09474-0 (cit. on pp. 18, 28, 29).

Baumuller, J., & Sopp, K. (2022). Double materiality and the shift from non-financial to European sustainability reporting: Review, outlook and implications. *JOURNAL*

OF APPLIED ACCOUNTING RESEARCH, *23*(1), 8–28. https://doi.org/10.1108/JAAR-04-2021-0114 (cit. on pp. 33, 34).

Baumuller, J., & Grbenic, S. O. (2021). Moving from Non-financial to Sustainability Reporting: Analyzing the EU Commission's Proposal for a Corporate Sustainability Reporting Directive (CSRD): Facta Universitatis, Series: Economics and Organization. *Facta Universitatis, Series: Economics and Organization*, *18*(4), 369–381 (cit. on p. 35).

Belenesi, M., Bogdan, V., & Popa, D. (2021). Disclosure Dynamics and Non-Financial Reporting Analysis. The Case of Romanian Listed Companies. *SUSTAINABILITY*, *13*(9). https://doi.org/10.3390/su13094732 (cit. on p. 23).

Beretta, V., Demartini, M., & Sotti, F. (2023). Board composition and textual attributes of non-financial disclosure in the banking sector: Evidence from the Italian setting after directive 2014/95/EU [publisher: Elsevier Ltd]. *Journal of Cleaner Production*, *385*. https://doi.org/10.1016/j.jclepro.2022.135561 (cit. on pp. 28, 29, 36).

Bigelli, M., Mengoli, S., & Sandri, S. (2023). Esg Score, Board Structure and the Impact of the Non-financial Reporting Directive on European Firms: Journal of Economics and Business. *Journal of Economics and Business*, *127*. https://doi.org/10.1016/j.jeconbus.2023.106133 (cit. on pp. 27, 31).

Bini, L., Schaper, S., Simoni, L., Giunta, F., & Nielsen, C. (2023). Mandatory non-financial disclosure: Is everybody on the same page about business model reporting? [publisher: Taylor and Francis Ltd.]. *Accounting Forum*, *47*(2), 198–222. https://doi.org/10.1080/01559982.2023.2170036 (cit. on pp. 12, 35, 37).

Bochenek, M. (2020). Analysis of the integrated reporting use in EU countries [publisher: Business Perspectives Ltd.]. *Problems and Perspectives in Management*, *18*(3), 106 (cit. on p. 22).

Borowicz, A., & Czerepko, J. (2023). Possible Consequences of Corporate Sustainability Reporting Directive on Polish Transport Companies [number-of-pages: 15 publisher-place: Warsaw publisher: University of Warsaw Centre for Europe Web of Science ID: WOS:001088103400006]. *STUDIA EUROPEJSKIE-STUDIES IN EUROPEAN AFFAIRS*, *27*(3), 105–119. https://doi.org/10.33067/SE.3.2023.6 (cit. on p. 23).

Breijer, R., & Orij, R. (2022). The Comparability of Non-Financial Information: An Exploration of the Impact of the Non-Financial Reporting Directive (NFRD, 2014/95/EU) [publisher: Routledge]. *Accounting in Europe*, *19*(2), 332–361. https://doi.org/10.1080/17449480.2022.2065645 (cit. on p. 20).

Broniewicz, E., Jastrzębska, E., & Lulewicz-Sas, A. (2024). Environmental disclosures according to ESRS in ESG reporting of selected banks in Poland [publisher: Wydawnictwo Ekonomia i Srodowisko]. *Economics and Environment*, *88*(1). https://doi.org/10.34659/eis.2024.88.1.719 (cit. on p. 34).

Caputo, F., Leopizzi, R., Pizzi, S., & Milone, V. (2020). The Non-Financial Reporting Harmonization in Europe: Evolutionary Pathways Related to the Transposition of the Directive 95/2014/EU within the Italian Context. *SUSTAINABILITY*, *12*(1). https://doi.org/10.3390/su12010092 (cit. on pp. 28, 31).

Carmo, C., & Ribeiro, C. (2022). Mandatory Non-Financial Information Disclosure under European Directive 95/2014/EU: Evidence from Portuguese Listed Companies. *SUSTAINABILITY*, *14*(8). https://doi.org/10.3390/su14084860 (cit. on p. 20).

Carungu, J., Di Pietra, R., & Molinari, M. (2021). Mandatory vs voluntary exercise on non-financial reporting: Does a normative/coercive isomorphism facilitate an increase in quality? *MEDITARI ACCOUNTANCY RESEARCH*, *29*(3), 449–476. https://doi.org/10.1108/MEDAR-08-2019-0540 (cit. on p. 35).

Černe, K., Hasić, T., & Galant, A. (2017). Implications of "Directive on Disclosure of Non-Financial and Diversity Information" on Croatian companies' reporting system. *Management International Conference*, 285–303 (cit. on pp. 21, 22, 24).

Chimonaki, C., Papadakisa, S., & Lemonakisa, C. (2024). Environmental, social, governance and gender diversity under the adoption of European Directive 2014/95/EU [[Online; accessed 2024-08-24]]. *Journal of Accounting & Management Information Systems/Contabilitate si Informatica de Gestiune*, *23*(1).

https://www.ceeol.com/search/article-detail?id=1228565. https://www.ceeol.com/search/article-detail?id=1228565 (cit. on p. 27).

Cordazzo, M., Bini, L., & Marzo, G. (2020). Does the EU Directive on non-financial information influence the value relevance of ESG disclosure? Italian evidence [number-of-pages: 14 publisher-place: Hoboken publisher: Wiley Web of Science ID: WOS:000549096400001]. *BUSINESS STRATEGY AND THE ENVIRONMENT*, *29*(8), 3470–3483. https://doi.org/10.1002/bse.2589 (cit. on pp. 12, 17).

Cuomo, F., Gaia, S., Girardone, C., & Piserà, S. (2024). The effects of the EU non-financial reporting directive on corporate social responsibility [publisher: Routledge]. *European Journal of Finance*, *30*(7), 726–752. https://doi.org/10.1080/1351847X.2022.2113812 (cit. on pp. 12, 20, 37).

De Cristofaro, T., & Gulluscio, C. (2023). In Search of Double Materiality in Non-Financial Reports: First Empirical Evidence. *SUSTAINABILITY*, *15*(2). https://doi.org/10.3390/su15020924 (cit. on pp. 33, 34).

Dencic-Mihajlov, K., & Spasic, D. (2016). Mandatory and Voluntary Disclosures of Serbian Listed Companies–Achieved Level and Some Recommendation for Improving Their Relevance: International Journal of Business and Economic Sciences Applied Research. *International Journal of Business and Economic Sciences Applied Research*, *9*(1), 27–38 (cit. on p. 36).

Dias, A., Rodrigues, L., & Craig, R. (2017). Corporate governance effects on social responsibility disclosures [publisher: University of Wollongong]. *Australasian Accounting, Business and Finance Journal*, *11*(2), 3–22. https://doi.org/10.14453/aabfj.v11i2.2 (cit. on pp. 20, 28, 29).

Dinh, T., Husmann, A., & Melloni, G. (2023). Corporate Sustainability Reporting in Europe: A Scoping Review. *ACCOUNTING IN EUROPE*, *20*(1), 91–119. https://doi.org/10.1080/17449480.2022.2149345 (cit. on p. 13).

Doni, F., Martini, S., Corvino, A., & Mazzoni, M. (2020). Voluntary versus mandatory non-financial disclosure EU Directive 95/2014 and sustainability reporting practices based on empirical evidence from Italy. *MEDITARI ACCOUNTANCY RESEARCH*, *28*(5), 781–802. https://doi.org/10.1108/MEDAR-12-2018-0423 (cit. on p. 17).

Dragija Kostic, M., Jovanovic, T., & Letica, M. (2017). Perspectives on and Obstacles to the Internal Reporting Reform at Higher Education Institutions–Case of Bosnia and Herzegovina, Croatia and Slovenia: Management. *Management*, *22*, 129–143 (cit. on p. 22).

Dragomir, V., Dumitru, M., Chersan, I., Gorgan, C., & Păunescu, M. (2024). Double Materiality Disclosure as an Emerging Practice: The Assessment Process, Impacts, Risks, and Opportunities [publisher: Routledge]. *Accounting in Europe*. https://doi.org/10.1080/17449480.2024.2339264 (cit. on pp. 23, 34).

Dumitru, V., Jinga, G., Stanila, O., & Dumitru, M. (2019). The impact of the European Directive 2014/95/EU on the energy companies' disclosures [issue: 1], *13*, 268–276. https://doi.org/10.2478/picbe-2019-0024 (cit. on pp. 13, 22, 23).

Esteban-Arrea, R., & Garcia-Torea, N. (2022). Strategic responses to sustainability reporting regulation and multiple stakeholder demands: An analysis of the Spanish EU non-financial reporting directive transposition [publisher: Emerald Group Holdings Ltd.]. *Sustainability Accounting, Management and Policy Journal*, *13*(3), 600–625. https://doi.org/10.1108/SAMPJ-07-2021-0292 (cit. on pp. 18, 38).

Eugénio, T., Gomes, S., Branco, M., & Morais, A. (2022). Non-Financial Reporting and Assurance: A New Opportunity for Auditors? Evidence from Portugal. *SUSTAINABILITY*, *14*(20). https://doi.org/10.3390/su142013469 (cit. on p. 20).

European Commission. (2014, October). Directive 2014/95/EU OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 22 October 2014, amending Directive 2013/34/EU as regards disclosure of non-financial and diversity information by certain large undertakings and groups (NFRD) [Doc ID: 32014L0095 Doc Sector: 3 Doc Title: Directive 2014/95/EU of the European Parliament and of the Council of 22 October 2014 amending Directive 2013/34/EU as regards disclosure of non-financial and diversity information by certain large undertakings and groups Text with EEA relevance Doc Type: L Usr_lan: en]. https://eur-lex.europa.eu/eli/dir/2014/95/oj (cit. on p. 12).

European Commission. (2017, July). Communication from the Commission — Guidelines on non-financial reporting (methodology for reporting non-financial information). https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52017XC0705(01) (cit. on pp. 32, 35).

European Commission. (2020). Summary report of the public consultation on the reviewof the non-financial reporting directive 20 February 2020 - 11 June 2020 [[Online; accessed 2024-07-30]]. https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12129-Revision-of-Non-Financial-Reporting-Directive/public-consultation_en (cit. on p. 32).

European Commission. (2022, December). Directive (EU) 2022/2464 of the European Parliament and of the Council of 14 December 2022 amending Regulation (EU) No 537/2014, Directive 2004/109/EC, Directive 2006/43/EC and Directive 2013/34/EU, as regards corporate sustainability reporting. [Usr_lan: EN]. https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX%3A32022L2464 (cit. on p. 32).

Fiandrino, S., Gromis di Trana, M., Tonelli, A., & Lucchese, A. (2022). The multi-faceted dimensions for the disclosure quality of non-financial information in revising directive 2014/95/EU [publisher: Emerald Group Holdings Ltd.]. *Journal of Applied Accounting Research*, *23*(1), 274–300. https://doi.org/10.1108/JAAR-04-2021-0118 (cit. on pp. 33, 36).

Fiechter, P., Hitz, J., & Lehmann, N. (2022). Real Effects of a Widespread CSR Reporting Mandate: Evidence from the European Union's CSR Directive [publisher: John Wiley and Sons Inc]. *Journal of Accounting Research*, *60*(4), 1499–1549. https://doi.org/10.1111/1475-679X.12424 (cit. on pp. 13, 38).

García-Benau, M., Bollas-Araya, H., & Sierra-García, L. (2022). Non-financial reporting in Spain. The effects of the adoption of the 2014 EU Directive. *REVISTA DE*

CONTABILIDAD-SPANISH ACCOUNTING REVIEW, *25*(1), 3–15. https://doi.org/10.6018/rcsar.392631 (cit. on p. 18).

García-Sánchez, I., Ortiz-Martínez, E., Marín-Hernández, S., & Aibar-Guzmán, B. (2023). How does the European Green Deal affect the disclosure of environmental information? *CORPORATE SOCIAL RESPONSIBILITY AND ENVIRONMENTAL MANAGEMENT*, *30*(6), 2766–2782. https://doi.org/10.1002/csr.2514 (cit. on pp. 17, 27–29, 31, 32, 38).

Gerwing, T., Kajüter, P., & Wirth, M. (2022). The role of sustainable corporate governance in mandatory sustainability reporting quality [publisher: Springer Science and Business Media Deutschland GmbH]. *Journal of Business Economics*, *92*(3), 517–555. https://doi.org/10.1007/s11573-022-01092-x (cit. on pp. 19, 28, 31, 33, 35).

Glaveli, N., Alexiou, M., Maragos, A., Daskalopoulou, A., & Voulgari, V. (2023). Assessing the Maturity of Sustainable Business Model and Strategy Reporting under the CSRD Shadow [publisher: Multidisciplinary Digital Publishing Institute (MDPI)]. *Journal of Risk and Financial Management*, *16*(10). https://doi.org/10.3390/jrfm16100445 (cit. on p. 35).

Gray, S. J. (1988). Towards a Theory of Cultural Influence on the Development of Accounting Systems Internationally. *Abacus*, *24*(1), 1–15. https://doi.org/10.1111/j.1467-6281.1988.tb00200.x (cit. on p. 13).

Grewal, J., Riedl, E. J., & Serafeim, G. (2019). Market Reaction to Mandatory Nonfinancial Disclosure: Management Science. *Management Science*, *65*(7), 3061–3084 (cit. on p. 12).

Habek, P., & Wolniak, R. (2016). Assessing the quality of corporate social responsibility reports: The case of reporting practices in selected European Union member states. *QUALITY & QUANTITY*, *50*(1), 399–420. https://doi.org/10.1007/s11135-014-0155-z (cit. on p. 21).

Hąbek, P. (2018). Are There Any Common Features of National Cultures of the Top Ten Sustainability Countries in European Union? *on European Integration 2018*, 380 (cit. on p. 25).

Hamed, R., Al-Shattarat, B., Al-Shattarat, W., & Hussainey, K. (2022). The impact of introducing new regulations on the quality of CSR reporting: Evidence from the UK. *JOURNAL OF INTERNATIONAL ACCOUNTING AUDITING AND TAXATION*, *46*. https://doi.org/10.1016/j.intaccaudtax.2021.100444 (cit. on pp. 19, 20, 27).

Hardies, K., Ohlrogge, F., Mentens, J., & Vandennieuwenhuysen, J. (2023). A Guide for Accounting Researchers to Conduct and Report Systematic Literature Reviews. *Behavioral Research in Accounting*, *36*, 1–23. https://doi.org/10.2308/BRIA-2022-042 (cit. on p. 15).

Hategan, C.-D., Pitorac, R.-I., & Milu, N.-D. (2021). Assessment of the mandatory non-financial reporting of romanian companies in the circular economy context [publisher: MDPI]. *International Journal of Environmental Research and Public Health*, *18*(24). https://doi.org/10.3390/ijerph182412899 (cit. on pp. 23, 38).

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*. https://doi.org/10.1016/j.jclepro.2023.138960 (cit. on pp. 19, 36).

Hofstede, G., Hofstede, G. J., & Minkov, M. (2014). *Cultures and Organizations: Software of the mind* [page: -]. Hakjisa. https://library.wur.nl/WebQuery/wurpubs/454134 (cit. on pp. 25, 27, 28).

Hofstede, G. (1980). Culture and Organizations [1664 citations (Semantic Scholar/DOI) [2024-07-24] 1295 citations (Crossref) [2024-07-24] publisher: Routledge]. *International Studies of Management & Organization*, *10*(4), 15–41. https://doi.org/10.1080/00208825.1980.11656300 (cit. on pp. 24, 25).

Hummel, K., & Jobst, D. (2024). An Overview of Corporate Sustainability Reporting Legislation in the European Union [publisher: Routledge]. *Accounting in Europe*. https://doi.org/10.1080/17449480.2024.2312145 (cit. on pp. 33, 38).

Hummel, K., & Rotzel, P. (2019). Mandating the Sustainability Disclosure in Annual Reports–Evidence from the United Kingdom: Schmalenbach Business Review. *Schmalenbach Business Review*, *71*(2), 205–247. https://doi.org/10.1007/s41464-019-00069-8 (cit. on pp. 19, 20, 38).

Janicka, M., & Sajnóg, A. (2023). Environmental information disclosure: A cross-country analysis from European Union public companies. *Ruch Prawniczy, Ekonomiczny i Socjologiczny*, *85*(3), 101–118 (cit. on pp. 22, 24).

Jovanovič, D., & Jovanović, N. (2022). Corporate GOVERNANCE CHALLENGES IN RELATION TO THE ESG REPORTING [publisher: University of Zagreb Faculty of Economics and Business]. *InterEULawEast*, *9*(2), 269–287. https://doi.org/10.22598/iele.2022.9.2.9 (cit. on p. 23).

Kaya, I. (2016). The Mandatory Social and Environmental Reporting: Evidence from France. In C. Zehir & E. Ozdemir (Eds.). https://doi.org/10.1016/j.sbspro.2016.07.130 (cit. on p. 18).

Khatri, I., & Kjærland, F. (2023). Sustainability reporting practices and environmental performance amongst nordic listed firms. *JOURNAL OF CLEANER PRODUCTION*, *418*. https://doi.org/10.1016/j.jclepro.2023.138172 (cit. on p. 21).

Kinderman, D. (2020). The Tenuous Link between CSR Performance and Support for Regulation: Business Associations and Nordic Regulatory Preferences regarding the Corporate Transparency Law 2014/95/EU: Business and Politics. *Business and Politics*, *22*(3), 413–448 (cit. on pp. 21, 22).

Klimczak, K. M., Hadro, D., & Meyer, M. (2023). Executive communication with stakeholders on sustainability: The case of Poland [number-of-pages: 23 publisher-place: Abingdon publisher: Routledge Journals, Taylor & Francis Ltd Web of Science ID: WOS:000999022500001]. *ACCOUNTING IN EUROPE*, *20*(3), 281–303. https://doi.org/10.1080/17449480.2023.2213242 (cit. on p. 23).

Korca, B., Costa, E., & Farneti, F. (2021). From voluntary to mandatory non-financial disclosure following Directive 2014/95/EU: An Italian case study. *ACCOUNTING IN EUROPE*, *18*(3), 353–377. https://doi.org/10.1080/17449480.2021.1933113 (cit. on pp. 12, 14, 17).

Korca, B. (2022). Sustainability Reporting in a Regulated Setting: Evidence from the European Union [publisher: Università degli studi di Trento]. https://iris.unitn.it/handle/11572/325568. https://iris.unitn.it/handle/11572/325568 (cit. on p. 35).

Krasodomska, J., Michalak, J., & Swietla, K. (2020). Directive 2014/95/EU Accountants' understanding and attitude towards mandatory non-financial disclosures in corporate reporting. *MEDITARI ACCOUNTANCY RESEARCH*, *28*(5), 751–779. https://doi.org/10.1108/MEDAR-06-2019-0504 (cit. on pp. 23, 24).

Krasodomska, J., Zarzycka, E., & Zieniuk, P. (2023). Voluntary sustainability reporting assurance in the European Union before the advent of the corporate sustainability reporting directive: The country and firm-level impact of Sustainable Development Goals [number-of-pages: 13 publisher-place: Hoboken publisher: Wiley Web of Science ID: WOS:001065025400001]. *SUSTAINABLE DEVELOPMENT*. https://www.webofscience.com/wos/woscc/summary/5bc06295-6024-48dc-9a8a-19b86c6c3521-e4f5e9c7/relevance/1. https://doi.org/10.1002/sd.2744 (cit. on p. 21).

Krasodomska, J., Zieniuk, P., & Kostrzewska, J. (2022). Reporting on Sustainable Development Goals in the European Union: What drives companies' decisions? [publisher: Emerald Publishing Limited]. *Competitiveness Review: An International Business Journal*, *33*(1), 120–146 (cit. on p. 21).

Kühn, A.-L., Stiglbauer, M., & Heel, J. (2014). Does mandatory CSR reporting lead to higher CSR transparency? The case of France [publisher: Virtus Interpress]. *Corporate Ownership and Control*, *11*(2 A), 29–45. https://doi.org/10.22495/cocv11i2p3 (cit. on pp. 18, 38).

La Torre, M., Sabelfeld, S., Blomkvist, M., & Dumay, J. (2020). Rebuilding trust: Sustainability and non-financial reporting and the European Union regulation. *MEDITARI ACCOUNTANCY RESEARCH*, *28*(5), 701–725. https://doi.org/10.1108/MEDAR-06-2020-0914 (cit. on pp. 32–34).

Lemos, K., Monteiro, S., Pereira, L., & Ribeiro, V. (2021). The Impact of Board of Director's Characteristics on Mandatory Non-Financial Reporting Compliance: A Study of Listed Companies in Portugal. In F. Bezzina (Ed.). https://doi.org/10.34190/MLG.21.060 (cit. on p. 28).

Lemos, K., Monteiro, S., & Marques, R. (2023, January). *Non-financial Disclosure of Large Portuguese Companies and Its Determinant Factors* [page: 102 DOI: 10.31410/ERAZ.S.P.2023.93]. (Cit. on p. 20).

Lin, W. (2022). Corporate Non-Financial Reporting in the UK: Diversions from the EU Sustainability Reporting Framework [number-of-pages: 18 publisher-place: Basel publisher: MDPI Web of Science ID: WOS:000840201600001]. *SUSTAINABILITY*, *14*(15), 9134. https://doi.org/10.3390/su14159134 (cit. on p. 20).

Lippai-Makra, E., Kovács, Z., & Kiss, G. (2022). The non-financial reporting practices of Hungarian listed public interest entities considering the 2014/95/EU Directive. *JOURNAL OF APPLIED ACCOUNTING RESEARCH*, *23*(1), 301–318. https://doi.org/10.1108/JAAR-04-2021-0086 (cit. on pp. 23, 24).

Lombardi, R., Cosentino, A., Sura, A., & Galeotti, M. (2022). The impact of the EU Directive on non-financial information: Novel features of the Italian case. *MEDITARI ACCOUNTANCY RESEARCH*, *30*(6), 1419–1448. https://doi.org/10.1108/MEDAR-06-2019-0507 (cit. on pp. 17, 33).

Lombardi, R., Schimperna, F., Paoloni, P., & Galeotti, M. (2022). The climate-related information in the changing EU directive on non-financial reporting and disclosure: First evidence by Italian large companies [publisher: Emerald Group Holdings Ltd.]. *Journal of Applied Accounting Research*, *23*(1), 250–273. https://doi.org/10.1108/JAAR-04-2021-0117 (cit. on p. 17).

Lopez, M. R., Ariza, L. R., Valentinetti, D., & Munoz, F. F. (2023). Risk Disclosures and Non-financial Reporting: Evidence in a New European Context: Scientific Annals of Economics and Business. *Scientific Annals of Economics and Business*, *70*(4), 547–565 (cit. on p. 33).

Lueg, K., Lueg, R., Andersen, K., & Dancianu, V. (2016). Integrated reporting with CSR practices: A pragmatic constructivist case study in a Danish cultural setting [publisher: Emerald Group Publishing Ltd.]. *Corporate Communications*, *21*(1), 20–35. https://doi.org/10.1108/CCIJ-08-2014-0053 (cit. on p. 22).

Matuszak, L., & Rózanska, E. (2021). Towards 2014/95/EU directive compliance: The case of Poland. *SUSTAINABILITY ACCOUNTING MANAGEMENT AND POLICY JOURNAL*, *12*(5), 1052–1076. https://doi.org/10.1108/SAMPJ-02-2020-0042 (cit. on p. 23).

Mazzotta, R., Bronzetti, G., & Veltri, S. (2020). Are mandatory non-financial disclosures credible? Evidence from Italian listed companies. *CORPORATE SOCIAL RESPONSIBILITY AND ENVIRONMENTAL MANAGEMENT*, *27*(4), 1900–1913. https://doi.org/10.1002/csr.1935 (cit. on pp. 33, 38).

Meidinger, E. (1987). Regulatory Culture: A Theoretical Outline*. *Law & Policy*, *9*(4), 355–386. https://doi.org/10.1111/j.1467-9930.1987.tb00416.x (cit. on p. 13).

Mezzanotte, F. (2023). Corporate sustainability reporting: Double materiality, impacts, and legal risk [publisher: Taylor and Francis Ltd.]. *Journal of Corporate Law Studies*. https://doi.org/10.1080/14735970.2024.2319058 (cit. on p. 33).

Michalak, J., Staszkiewicz, P., & Waniak-Michalak, H. (2023). Is academic recommendation translated into the European Union corporate sustainability reporting directive proposal? [publisher: Elsevier Ltd]. *Journal of Cleaner Production*, *412*. https://doi.org/10.1016/j.jclepro.2023.137186 (cit. on p. 13).

Miettinen, M. (2024). Are materiality determination practices evolving in the wake of increasing legislation on sustainability reporting? Findings from EU pharmaceutical companies' reports [publisher: Emerald Publishing]. *International Journal of Law* *and Management*, *66*(3), 363–392. https://doi.org/10.1108/IJLMA-09-2023-0221 (cit. on pp. 33, 34).

Mio, C., Fasan, M., Marcon, C., & Panfilo, S. (2020). The predictive ability of legitimacy and agency theory after the implementation of the EU directive on non-financial information. *CORPORATE SOCIAL RESPONSIBILITY AND ENVIRONMENTAL MANAGEMENT*, *27*(6), 2465–2476. https://doi.org/10.1002/csr.1968 (cit. on pp. 28, 29).

Mion, G., & Adaui, C. (2019). Mandatory Nonfinancial Disclosure and Its Consequences on the Sustainability Reporting Quality of Italian and German Companies. *SUSTAINABILITY*, *11*(17). https://doi.org/10.3390/su11174612 (cit. on pp. 19, 38).

Mittelbach-Hörmanseder, S., Hummel, K., & Rammerstorfer, M. (2021). The information content of corporate social responsibility disclosure in Europe: An institutional perspective. *EUROPEAN ACCOUNTING REVIEW*, *30*(2), 309–348. https://doi.org/10.1080/09638180.2020.1763818 (cit. on p. 25).

Moggi, S., Lehman, G., & Pagani, A. (2023). The juridification of social accounting and the transposition process of the non-financial reporting directive 2014/95/EU [publisher: Emerald Publishing]. *Meditari Accountancy Research*, *31*(7), 185–211. https://doi.org/10.1108/MEDAR-01-2023-1897 (cit. on p. 12).

Nicolo, G., Zanellato, G., Manes-Rossi, F., & Tiron-Tudor, A. (2021). Corporate reporting metamorphosis: Empirical findings from state-owned enterprises [46 citations (Semantic Scholar/DOI) [2024-07-24] 38 citations (Crossref) [2024-07-24] publisher: Routledge _eprint: https://doi.org/10.1080/09540962.2020.1719633]. *Public Money & Management*, *41*(2), 138–147. https://doi.org/10.1080/09540962.2020.1719633 (cit. on p. 35).

Nicolò, G., Zampone, G., Sannino, G., & De Iorio, S. (2022). Sustainable corporate governance and non-financial disclosure in Europe: Does the gender diversity matter? *JOURNAL OF APPLIED ACCOUNTING RESEARCH*, *23*(1), 227–249. https://doi.org/10.1108/JAAR-04-2021-0100 (cit. on pp. 27, 29).

Panfilo, S., & Krasodomska, J. (2022). Climate Change Risk Disclosure in Europe: The Role of Cultural-Cognitive, Regulative, and Normative Factors. *ACCOUNTING IN EUROPE*, *19*(1), 226–253. https://doi.org/10.1080/17449480.2022.2026000 (cit. on p. 25).

Papa, M., Carrassi, M., Muserra, A., & Wieczorek-Kosmala, M. (2022). The impact of the EU nonfinancial information directive on environmental disclosure: Evidence from Italian environmentally sensitive industries. *MEDITARI ACCOUNTANCY RESEARCH*, *30*(7), 87–120. https://doi.org/10.1108/MEDAR-03-2021-1247 (cit. on pp. 12, 17).

Petruzzelli, S., & Badia, F. (2024). The quality assessment of stakeholder engagement disclosure in the EU mandatory non-financial reporting framework [publisher: Emerald Publishing]. *Journal of Applied Accounting Research*, *25*(1), 126–148. https://doi.org/10.1108/JAAR-11-2022-0290 (cit. on p. 12).

Pizzi, S. (2024). Regulating Sustainability Reporting in Europe: De Jure Harmonisation or De Facto Standardisation? [publisher: Routledge]. *Accounting in Europe*. https://doi.org/10.1080/17449480.2024.2330976 (cit. on p. 17).

Pizzi, S., Del Baldo, M., Caputo, F., & Venturelli, A. (2022). Voluntary disclosure of Sustainable Development Goals in mandatory non-financial reports: The moderating role of cultural dimension. *JOURNAL OF INTERNATIONAL FINANCIAL MANAGEMENT & ACCOUNTING*, *33*(1), 83–106. https://doi.org/10.1111/jifm.12139 (cit. on pp. 12, 13, 21, 25, 35–37).

Pizzi, S., Principale, S., Fasiello, R., & Imperiale, F. (2023). The institutionalisation of social and environmental accounting practices in Europe. *JOURNAL OF APPLIED ACCOUNTING RESEARCH*, *24*(5), 816–838. https://doi.org/10.1108/JAAR-07-2022-0190 (cit. on pp. 27, 29).

Pizzi, S., Venturelli, A., & Caputo, F. (2021). The "comply-or-explain" principle in directive 95/2014/EU. A rhetorical analysis of Italian PIEs [publisher: Emerald Group Holdings Ltd.]. *Sustainability Accounting, Management and Policy Journal*, *12*(1), 30–50. https://doi.org/10.1108/SAMPJ-07-2019-0254 (cit. on p. 12).

Posadas, S. C., Ruiz-Blanco, S., Fernandez-Feijoo, B., & Tarquinio, L. (2023). Institutional isomorphism under the test of Non-financial Reporting Directive. Evidence from Italy and Spain [number-of-pages: 23 publisher-place: Bingley publisher: Emerald Group Publishing Ltd Web of Science ID: WOS:000912784400001]. *MEDITARI ACCOUNTANCY RESEARCH*, *31*(7), 26–48. https://doi.org/10.1108/MEDAR-02-2022-1606 (cit. on pp. 13, 16–18).

Primec, A., & Belak, J. (2022). Sustainable CSR: Legal and Managerial Demands of the New EU Legislation (CSRD) for the Future Corporate Governance Practices [publisher: MDPI]. *Sustainability (Switzerland)*, *14*(24). https://doi.org/10.3390/su142416648 (cit. on pp. 23, 24, 35, 36).

Radu, O. M., Dragomir, V. D., & Hao, N. (2023). Company-Level Factors of Non-Financial Reporting Quality under a Mandatory Regime: A Systematic Review of Empirical Evidence in the European Union [number-of-pages: 32 publisher-place: Basel publisher: MDPI Web of Science ID: WOS:001116180200001]. *SUSTAINABILITY*, *15*(23), 16265. https://doi.org/10.3390/su152316265 (cit. on pp. 27–29).

Raith, D. (2023). The contest for materiality. What counts as CSR? [publisher: Emerald Publishing]. *Journal of Applied Accounting Research*, *24*(1), 134–148. https://doi.org/10.1108/JAAR-04-2022-0093 (cit. on pp. 19, 32).

Raucci, D., & Tarquinio, L. (2020). Sustainability Performance Indicators and Non-Financial Information Reporting. Evidence from the Italian Case. *ADMINISTRATIVE SCIENCES*, *10*(1). https://doi.org/10.3390/admsci10010013 (cit. on pp. 17, 38).

Rep, A., Malis, S. S., & Dumancic, K. (2022). Factors Influencing the Disclosure of Additional Financial and Non-financial Information by Large Enterprises: Management. *Management*, *27*(1), 287–308 (cit. on p. 29).

Rezaee, Z., Homayoun, S., Poursoleyman, E., & Rezaee, N. J. (2023). Comparative Analysis of Environmental, Social, and Governance Disclosures: Global Finance Journal.

Global Finance Journal, *55*. https://doi.org/10.1016/j.gfj.2022.100804 (cit. on p. 31).

Ruiz, S., Romero, S., & Fernandez-Feijoo, B. (2021). Stakeholder engagement is evolving: Do investors play a main role? [publisher: John Wiley and Sons Ltd]. *Business Strategy and the Environment*, *30*(2), 1105–1120. https://doi.org/10.1002/bse.2674 (cit. on pp. 33, 34).

Scandurra, G., & Thomas, A. (2023). The SDGs and Non-Financial Disclosures of Energy Companies: The Italian Experience. *SUSTAINABILITY*, *15*(17). https://doi.org/10.3390/su151712882 (cit. on p. 35).

Schröder, P. (2022). Mandatory non-financial reporting in the banking industry: Assessing reporting quality and determinants [publisher: Cogent OA]. *Cogent Business and Management*, *9*(1). https://doi.org/10.1080/23311975.2022.2073628 (cit. on pp. 19, 22).

Sierra-Garcia, L., Garcia-Benau, M., & Bollas-Araya, H. (2018). Empirical analysis of non-financial reporting by spanish companies [publisher: MDPI]. *Administrative Sciences*, *8*(3). https://doi.org/10.3390/admsci8030029 (cit. on p. 18).

Simoni, L., Schaper, S., & Nielsen, C. (2022). Business Model Disclosures, Market Values, and Earnings Persistence: Evidence From the UK. *ABACUS-A JOURNAL OF ACCOUNTING FINANCE AND BUSINESS STUDIES*, *58*(1), 142–173. https://doi.org/10.1111/abac.12233 (cit. on pp. 35, 37).

Singhania, M., & Saini, N. (2022). Quantification of ESG Regulations: A Cross-Country Benchmarking Analysis [number-of-pages: 9 publisher-place: New Delhi publisher: SAGE Publications India Pvt Ltd Web of Science ID: WOS:000800314000003]. *VISION-THE JOURNAL OF BUSINESS PERSPECTIVE*, *26*(2), 163–171. https://doi.org/10.1177/09722629211054173 (cit. on p. 21).

Szabados, T. (2021). Multilevel hardening in progress – Transition from soft towards hard regulation of CSR in the EU [publisher: SAGE Publications Ltd]. *Maastricht Journal of European and Comparative Law*, *28*(1), 83–101. https://doi.org/10.1177/1023263X20958379 (cit. on p. 12).

Tamm, P., & Gurvits-Suits, N. A. (2023). Development of Non-financial Reporting: The Case of Estonian Listed Companies [number-of-pages: 11 publisher-place: Kaunas publisher: Kauno Technologijos Univ Web of Science ID: WOS:001072056500015]. *EUROPEAN INTEGRATION STUDIES*, (17), 199–209. https://doi.org/10.5755/j01.eis.1.17.34060 (cit. on p. 23).

Tarquinio, L., Posadas, S., & Pedicone, D. (2020). Scoring Nonfinancial Information Reporting in Italian Listed Companies: A Comparison of before and after the Legislative Decree 254/2016. *SUSTAINABILITY*, *12*(10). https://doi.org/10.3390/su12104158 (cit. on p. 17).

Tocev, T., Malchev, B., Trpeska, M., & Minovski, Z. (2022). How DOES EU MEMBERSHIP AFFECT NON-FINANCIAL REPORTING? EVIDENCE FROM SELECTED BALKAN COUNTRIES [publisher: Faculty of Economics-Skopje, Ss. Cyril and Methodius University in Skopje]. https://repository.ukim.mk/handle/20.500.12188/24435. https://repository.ukim.mk/handle/20.500.12188/24435 (cit. on p. 23).

Tozser, D., Lakner, Z., Sudibyo, N., & Boros, A. (2024). Disclosure Compliance with Different ESG Reporting Guidelines: The Sustainability Ranking of Selected European and Hungarian Banks in the Socio-Economic Crisis Period. *ADMINISTRATIVE SCIENCES*, *14*(3). https://doi.org/10.3390/admsci14030058 (cit. on p. 22).

Traxler, A., Greiling, D., Freinbichler, M., & Mayerhofer, P. (2023). Does the Directive of the European Union on disclosure of non-financial and diversity information improve reporting? *JOURNAL OF ACCOUNTING AND ORGANIZATIONAL CHANGE*, *19*(6), 70–90. https://doi.org/10.1108/JAOC-04-2022-0071 (cit. on p. 12).

Velte, P. (2023a). Does sustainable board governance drive corporate social responsibility? A structured literature review on European archival research [publisher: Emerald Publishing]. *Journal of Global Responsibility*, *14*(1), 46–88. https://doi.org/10.1108/JGR-05-2022-0044 (cit. on pp. 25, 27).

Velte, P. (2023b). Sustainable board governance and environmental performance: European evidence. *BUSINESS STRATEGY AND THE ENVIRONMENT*. https://doi.org/10.1002/bse.3654 (cit. on p. 27).

Venturelli, A., Caputo, F., Cosma, S., Leopizzi, R., & Pizzi, S. (2017). Directive 2014/95/EU: Are Italian companies already compliant? [publisher: MDPI]. *Sustainability (Switzerland)*, *9*(8). https://doi.org/10.3390/su9081385 (cit. on p. 17).

Venturelli, A., Caputo, F., Leopizzi, R., & Pizzi, S. (2019). The state of art of corporate social disclosure before the introduction of non-financial reporting directive: A cross country analysis [publisher: Emerald Group Holdings Ltd.]. *Social Responsibility Journal*, *15*(4), 409–423. https://doi.org/10.1108/SRJ-12-2017-0275 (cit. on pp. 16, 19, 20).

Zanellato, G., & Tiron-Tudor, A. (2022). How cultural dimensions are shaping social expectations: The case of European state-owned enterprises' nonfinancial reporting [publisher: Emerald Publishing Limited]. *Journal of Applied Accounting Research*, *23*(1), 99–121 (cit. on pp. 21, 25).

Chapter 2

Precision and Perception: Analyzing the Credibility in ESG Disclosures

Melanie Luxembourger* / Maxime Pettinger†

University of Luxembourg /STATEC, Luxembourg

Abstract

This research investigates the effect of time-horizon precision on the perceived credibility of Environmental, Social, and Governance (ESG) disclosures. We examine whether clear and specific timelines in disclosures can enhance confidence in the report and identify specific groups of individuals for whom this precision is most impactful. Using a quasi-experimental design within a survey, participants were randomly assigned to one of two groups: a treatment group that received an ESG text fragment with precise time horizon details, and a control group for which the text fragment did not specify a forecast horizon. The established Perceived Credibility (PERCRED) scale is employed to capture participants' assessments of the disclosures. Our findings reveal that the effectiveness of time-horizon precision is contingent upon the reader's familiarity with ESG or firm reporting, as significant positive effects were observed among subgroups familiar with ESG or firm reporting. This research contributes to the discourse on ESG reporting by demonstrating the role of precision in influencing readers perceptions and the strategic importance of transparent reporting practices.

Keywords—ESG reporting, Credibility Perception, Time-Horizon Precision, Sustainability Disclosure

1 Introduction

Despite increasing investor demand for ESG data, concerns about the credibility of ESG disclosures persist, as they are often viewed as promotional tools rather than reliable sources of information (Amel-Zadeh and Serafeim, 2018; Dando and Swift, 2003; van Duuren et al., 2016).

This research paper investigates how time-horizon precision in ESG reporting influences stakeholders' trust and the perceived credibility of the information. By focusing on how precision influences stakeholders' trust, this study examines whether clear and specific timelines in ESG disclosures can enhance confidence in the reports, affecting organizational reputation and thereby potentially investment decisions. Additionally, we examine the factors explaining discrepancies in perceived credibility and identify specific groups of individuals for which precision is particularly effective in enhancing the perceived credibility.

The research question inherently calls for an experimental design, as we seek to determine the causal impact of varying levels of time-horizon precision on the perceived credibility of ESG disclosures. Our methodology includes an analysis of perceived credibility by readers of ESG text fragments presented with different levels of time horizon preciseness. We use the Perceived Credibility (PERCRED) scale developed by Lock and Seele, 2017 to capture the multi-dimensional construct of credibility in the context of ESG communication. Post-experiment analysis showed no significant differences in covariates between the control and treatment groups, suggesting that the randomization process was successful in balancing covariates.

Our investigation reveals that the treatment's impact is diluted when considering the entire population, making it non-significant. However, when focusing on specific sub-populations, such as those with prior experience reading firm reports or familiarity with ESG, the treatment shows a significant positive impact. This suggests that the treatment's effectiveness is contingent upon the recipient's background and familiarity with related content. The significant impact observed among those who had read a firm report can be intuitively explained by their prior experience, which provides them with a point of comparison. This prior exposure likely enables them to better discern and appreciate the details in the treatment, leading to a more pronounced effect.

These findings underscore the strategic importance of precise ESG disclosures as a mean to foster trust, promote transparency, and increase corporate legitimacy in the eyes of stakeholders. By demonstrating the direct link between the precision of disclosures and stakeholder perceptions, our findings emphasize the imperative for organizations to adopt precise and transparent reporting practices. These practices serve as a cornerstone of responsible business conduct and sustainable development, driving positive outcomes across the board.

The remainder of the paper is structured as follows: Section 2 reviews the related literature and develops the hypotheses, focusing on the credibility of ESG disclosures, the role of precision, and the joint effect of experience. Section 3 outlines the methodology, including the experimental design, treatment and outcome variables, and the models used for analysis. Section 4 presents the data and assesses the balance of covariates to ensure the validity of the experimental results. Section 5 discusses the results, including hypothesis testing and participant feedback. Finally, Section 6 concludes with the key findings, their implications, and recommendations for future research and practice.

2 Related Literature and Hypothesis Development

2.1 Credibility

Increasing attention to sustain ability has driven investor demand for ESG data (Amel-Zadeh and Serafeim, 2018), making accurate and reliable ESG information essential for informed investment decisions (van Duuren et al., 2016). However, despite ongoing efforts to improve corporate transparency, stakeholders historically have had difficulty placing their trust in the legitimacy of ESG disclosures, with a widespread skepticism finding that ESG reports are often perceived as promotional tools rather than dependable sources, largely due to concerns about the authenticity and reliability of the data provided (Dando and Swift, 2003).

As such, increasing credibility in ESG reporting is related to building trust among stakeholders and strengthening the firm's legitimacy in the public eye. When a firm's ESG disclosure is perceived as credible, it is more likely to trigger stakeholders to view the firm as legitimate in both a pragmatic sense (believing the firm's actions align with their own interests) and a cognitive sense (believing the firm's actions are in line with societal norms and values) (Lock and Schulz-Knappe, 2019). In other words, if stakeholders trust the ESG information, they are more likely to believe that the company is both acting in ways that benefit them and that are socially acceptable, enhancing the overall perception of the company's legitimacy.

The credibility of reports is a multidimensional perception construct that encompasses several dimensions (Balluchi et al., 2021; Lock and Seele, 2017), and is influenced by various factors. Such factors include firm-level and contextual factors (Lock and Seele, 2016), such as historical performance and reputation of the firms (Athanasakou and Hussainey, 2014; Torelli et al., 2020). Other factors influencing the credibility of the reports lie in the implementation details, such as the clarity and accessibility of the information, studied in research in terms of readability and formulation of the information (Boukes and LaMarre, 2021; Hoozée et al., 2019). Other important factors influencing credibility are the relevance and significance of the content of the report, studied as the materiality of the subjects covered (Reimsbach et al., 2018).Finally the transparency and verification of the information can heavily influence the credibility of the report, studied in research as the reference explicitness and assurance of the information (Baier et al., 2021; Reimsbach et al., 2020; Tyson and Adams, 2020). In our present study, we examine the influence of the precision of the information in the report on its credibility, conceptualized through time-horizon precision.

2.2 Precision

Companies can effectively fulfill their accountability responsibilities only when their reports are comprehensive, ensuring an accurate and impartial evaluation of ESG information that is relevant to their stakeholders (Adams, 2004). To be considered 'comprehensive,' published ESG materials must go beyond merely stating intentions; they should also provide detailed insights into related activities and progress towards stated objectives (Bouten et al., 2011). This is what we consider to be forward-looking information.

Forward-looking information, often included in management discussion and analysis sections, typically reflects shorter-term financial or operational expectations that stakeholders can assess in subsequent periods (Muslu et al., 2015). However, the tone of forward-looking statements can be biased, potentially affecting stakeholder perceptions (Schleicher and Walker, 2010). Indeed, companies often hedge their forward-looking statements using specific linguistic techniques that reduce the perceived certainty of future events to manage expectations and protect themselves legally (McLaren-Hankin, 2008). In this context, precision becomes a critical tool for enhancing the credibility of such statements.

Although forward-looking disclosures add value across all types of corporate reports, time horizon precision holds unique significance for ESG disclosures. Indeed, while ESG disclosures provide valuable insights into a company's sustainability efforts, they inherently carry a degree of uncertainty due to the extended timeframes required to observe the impact of sustainable projects (Carney, 2015)[1]. Sustainable goals typically require longer-term commitments—such as carbon neutrality by 2050 or substantial emission reductions by 2030—which extend the period over which stakeholders can observe results. This temporal gap poses a challenge to the credibility of ESG information as this extended timeline increases the inherent uncertainty of ESG outcomes, which can undermine stakeholders' trust.

However, when companies establish precise milestones within their ESG disclosures, they

[1]This is termed by Carney, 2015 as "the tragedy of the horizon"; a paradox in which market players must take immediate action to address long-term risks but do not have a short-term observed incentive.

provide stakeholders with a more concrete framework for tracking accountability over the long term. Therefore, while time horizon information is beneficial in general corporate disclosures, its precision is particularly critical in ESG reporting to confirm a company's commitment to its sustainability goals.

This suggests that when companies present ESG information with vague or indefinite timelines, it may lead to skepticism among stakeholders regarding the firm's commitment to its sustainability goals and its overall legitimacy. Conversely, when companies provide more precise time horizons in their ESG disclosures, it can signal a higher level of confidence and accountability. This leads to the following hypothesis:

Hypothesis 0: The Perceived Credibility is higher when the time-horizon of ESG disclosures is more precise.

While this hypothesis expects a direct effect, further investigation of the literature suggests that the actual impact of precision might be contingent on specific reader characteristics.

2.3 Joint effect of experience

Behavioral economics research in accounting indicates that individual behavior is shaped not only by economic incentives and available information but also by personal preferences, abilities, experiences, and other characteristics (Hanlon et al., 2022). Consequently, the evaluation of ESG disclosures is influenced by subjective judgments, which evolve as individuals acquire more knowledge and experience, significantly shaping perceptions of credibility (Tan et al., 2015).

Research suggests that financial professionals, through their domain-specific knowledge and experience, have refined cognitive abilities enabling them to outperform non-professionals in judgment tasks (Frederick and Libby, 1986; Libby and Luft, 1993). Experts' superior performance is attributed to their developed mental models and integration patterns, which allow them to effectively recognize and assess relevant information (Bonner, 2008; Libby and Frederick, 1990). Consequently, it has been found that experienced financial professionals are more adept than non-professionals at discerning and interpreting crucial information (Bonner and Lewis, 1990; Libby and Luft, 1993) and as such will exhibit greater sensitivity than non-professionals to relevant information in assessing firm performance (Henry and Peytcheva, 2020).

Moreover, report readers can have different specialization which can affect the perception of ESG information: specialized readers, who have extensive expertise in specific ESG topics, and generalist readers, who have a broader range of interests (Crilly et al., 2016; Hoozée et al., 2019). Experts, with their specialized knowledge, are adept at discerning the authenticity of companies' ESG practices, effectively distinguishing between implementers, who genuinely execute their ESG principles, and decouplers, who fail to do so (Crilly et al., 2016). Their skill stems from their extensive knowledge of ESG themes and processes, allowing them to conduct a thorough examination of ESG disclosures. In contrast, generalist readers often approach ESG disclosures with skepticism, more susceptible to perceiving ESG efforts as greenwashing, a practice where companies exaggerate or fabricate their engagement to environmental and social responsibilities (Torelli et al., 2020). This skepticism among non-experts emphasizes the need for transparency and clarity in ESG reporting, as the perceived credibility of these disclosures can significantly impact stakeholders' trust and confidence in a company's ESG commitments.

Given that 'expertise' with reporting is an inherently abstract concept that varies significantly among readers, we consider 'experts' here as readers who possess varying levels of experience with reports, and thus have a degree of familiarity with reporting practices. In this way, we operationalize "expertise" not as a fixed attribute but as a continuum of experience, ranging from basic familiarity to advanced professional involvement. This approach allows us to capture a broad range of reader capacities, from those who have occasionally engaged with reports to those who regularly analyse them in depth.

Taking into account this conceptualization of expertise, and its relationship with experience and familiarity, we expect the positive impact of precise time-horizons on the perceived credibility of ESG disclosures to be significant among individuals with expertise or familiarity with ESG or corporate reporting, and as such refine the previously stated hypothesis, H0, as:

H1: *Precise time-horizons have a positive effect on the perceived credibility of disclosures within target populations.*

To further refine, we propose that the impact of precise time-horizons on the perceived credibility of ESG disclosures is null for individuals without prior experience in reading reports, and significantly greater for those with expertise or familiarity in ESG or corporate reporting, such that:

H2: *For individuals with limited or no experience in ESG or corporate reporting, precise time-horizons do not significantly impact perceived disclosure credibility.*

These hypotheses underscore the idea that the perception of ESG disclosures is not only determined by the information presented but is significantly affected by the reader's background. A specialized reader with extensive experience in analyzing reports is likely to have a different perception of the same disclosure compared to a generalist reader with limited ESG exposure and who has never read or performed deep analysis on firm reports. These factors— levels of familiarity with reporting practices —act as lenses through which ESG disclosures are interpreted, influencing the perceived credibility of the information based on the depth of understanding and critical engagement with ESG content.

3 Methodology

Given the research question, a simple between-subjects experimental design is particularly relevant to determine the causal impact of varying levels of time-horizon precision on the perceived credibility of ESG disclosures.

3.1 Experimental design and treatment effect

3.1.1 Treatment Variable

The study was conducted through an online survey, where participants were all presented an ESG text fragment. There were two versions of this text based on time-horizon precision: one where ESG disclosures do not specify a forecast horizon, lacking any reference to time, and one where ESG disclosures include precise time-horizons and reference a specific point in time. Participants were randomly assigned to either the treatment or the control group, meaning they received either the precise version of the text fragment (treatment group) or the vague version (control group).

This random assignment is critical because it ensures that the treatment is applied independently of any other characteristics that could influence the outcome, thereby mitigating the risk of selection bias and omitted variable bias, and even theoretically eliminating them altogether.

All text fragments conveyed identical information, with the only difference being the level of time-horizon precision. The content of these fragments was adapted from the 2022 annual report of a real-life company, anonymized as *ABC*.

3.1.2 Recruitment Process

The sample comprised two main types of participants: novices (students) and experienced professionals.

To recruit student participants, a direct approach was employed by visiting various Bachelor's and Master's classes in disciplines related to Accounting, Management, and Economics. These visits were strategically scheduled towards the end of class sessions, allowing a portion of class time to be dedicated to explaining the survey and encouraging voluntary participation. Ethical considerations were emphasized by ensuring that students were informed of their right to opt out without any pressure or obligation to complete the survey. This recruitment effort resulted in a final sample of 115 students, including

97 generalists and 18 ESG-oriented participants.

For the recruitment of professional participants, the survey was promoted and distributed via LinkedIn, targeting specific groups and individuals aligned with the study's demographic requirements, specifically ESG-oriented professionals. The goal was to ensure adequate representation of ESG experts, whose insights are particularly valuable to this research, alongside generalist professionals. The survey link was accompanied by a description outlining the study's purpose and the anticipated time commitment. To maximize outreach, participants were also encouraged to share the survey with colleagues and peers who might be relevant to the study, thereby leveraging the network effect. This effort yielded a final sample of 106 professionals, with 47 ESG-oriented and 59 generalist participants.

While many previous experimental studies primarily utilize student participants due to the ease of recruitment (e.g. Baier et al., 2021; Hoozée et al., 2019), this study distinguishes itself by incorporating a substantial sample of actual professionals, thus enhancing the robustness and applicability of the findings. In total, the sample consists of 221 participants, combining both students and professionals.

3.1.3 Outcome Variable: Measurement of Credibility

To measure the perceived credibility of the ESG report, we rely on the PERCRED scale developed by Lock and Seele, 2017. This scale is designed to measure the perceived credibility of ESG communications across four dimensions: truth, sincerity, appropriateness, and understandability. Each dimension is assessed through a set of statements that participants rated using a 7-point Likert scale. The response options ranged from 1 (strongly disagree) to 7 (strongly agree), allowing participants to express their level of agreement with each statement. A detailed overview of the statements and the scale used is provided in **Appendix B**.

The scale captures the multi-dimensional construct of credibility through the evaluation of the following dimensions:

- *Understandability:* Participants assess how clear and easy to comprehend the information in the disclosure is.
- *Truth:* Participants assess how factually accurate and reliable the information seems.
- *Sincerity:* Participants judge whether the company appears honest and genuine in its communications.
- *Appropriateness:* Participants evaluate how relevant and contextually suitable the information is within the company's social and industrial environment.

Responses for each dimension were aggregated to create individual factor scores, which were then combined into a single PERCRED score representing the overall perceived credibility. This process follows the sequential approach outlined by Lock and Seele, 2017, using the weights and factor loadings they provided. Specifically, the values for the four dimensions (truth, sincerity, appropriateness, and understandability) were calculated by linearly combining the scores from 16 Likert-scale items, which were then aggregated into the overall PERCRED score for each respondent.

We decide to rely on this measure as it was developed through a rigorous process following the best practices described for instance by Boateng et al., 2018, and has previously demonstrated a high reliability score (Lock and Seele, 2017). To assess the reliability of the scale in our study, we employ both Cronbach's alpha and Guttman's lambda 6. While Cronbach's alpha is commonly used in applied research, it has been shown to suffer from significant limitations, leading to recommendations for the use of Guttman's lambdas as more reliable alternatives (see e.g.: Sijtsma, 2009). As such, we use both measures to ensure a comprehensive assessment of reliability.

3.2 Models and analysis

3.2.1 Estimation strategy

In this study, we define expertise as familiarity with ESG or corporate reporting, measured by prior experience in reading, analyzing, or working with such reports. While expertise in a broader sense includes not only experience but also deeper knowledge and critical evaluation skills, for the purposes of our analysis, we use experience as a practical proxy for expertise. This allows us to distinguish between individuals with varying levels of engagement with ESG disclosures, recognizing that repeated exposure and practical involvement with reporting can enhance one's ability to assess the credibility of precise time-horizons.

Our core hypothesis posits that the treatment—providing precise temporal details in ESG reports—enhances the perceived credibility of these commitments, particularly among individuals with relevant expertise. There are two primary methods to account for this conditional effect when estimating the treatment impact: separate-sample and pooled-sample estimation of interactive effects.

The separate-sample approach involves estimating the treatment effect within the subset of individuals identified as experts. We use this method to focus exclusively on the population that is presumed to be influenced by the treatment, allowing for a direct estimation of the Average Treatment Effect (ATE) for this group. The separate-sample model estimated is the following:

$$Y_S = \alpha + \beta\delta_S + \epsilon_S \quad \text{where} \quad S = \{i \mid X_i = 1\}, \# (1)$$

where δ is a binary variable indicating whether the individual S has been treated, and β is the average treatment effect (ATE) for the population of 'experts', and X is a binary variable indicating whether the individual i possesses an expertise in, or at least is familiar with, ESG or firm reporting.

On the other hand, the pooled-sample approach combines data from both experts and non-experts, estimating the treatment effect across the entire sample while simultaneously testing the interaction between treatment and expertise. By including an interaction term between treatment and expertise, the pooled-sample approach can quantify the differential treatment effect for experts compared to non-experts. The estimated pooled-sample model takes the following form:

$$Y_i = \alpha + \beta_X X_i + \beta_\delta \delta_i + \beta_{\mathrm{X}\delta} \left(X_i \times \delta_i\right) + \epsilon_i, \quad \#\,(2)$$

where β_X and β_δ denote the main effects of the 'expertise' and the treatment binary variables, respectively, while the interaction effect is given by $\beta_{\mathrm{X}\delta}$.

If $\widehat{\beta}_{X\delta}$, the estimate of $\beta_{\mathrm{X}\delta}$, is different from zero, we can conclude that the impact of the treatment is conditioned on the expertise of the respondent. The value of $\widehat{\beta}_{X\delta}$ gives the estimated magnitude of the between-group difference in treatment effect. The estimate of the treatment effect on the population of 'experts', β, is obtained by taking the sum of the main effect of the treatment and the interaction effect, i.e.:

$$\widehat{\beta} = \widehat{\beta}_\delta + \widehat{\beta}_{\delta X}.\#\,(3)$$

While both methods are valid for estimating the conditional effects of the treatment, they are not equivalent. The pooled-sample approach is more efficient in terms of using all the available data, which is a necessary condition for testing **H2**, which allows to test a full set of interactive hypotheses (Franzese and Kam, 2009), providing a straightforward assessment of the difference in treatment effect between experts and non-experts, and therefore allowing to straightforwardly test **H2**. In contrast, the separate-sample approach is better suited for **H1**. By isolating the expert group, this approach not only provides direct estimates of the treatment effect and inferential measures but also avoids the inclusion of interaction terms, which could possibly induce collinearity and lead to a loss of efficiency[2].

3.2.2 Assessing the balance of covariates

In any randomized experiment, ensuring the balance of covariates across treatment and control groups is crucial to attribute any observed effect to the treatment rather than to pre-existing differences between groups. At the population level, random assignment theoretically ensures that both groups should have similar characteristics on average, thereby reducing the risk of bias due to over or under representation of confounding characteristics in one or the other group. This implies that statistical tests of balance heterogeneity, being inferential by nature, are unnecessary and potentially misleading when applied to experimental data with random assignment of the treatment (Senn, 1994)[3]. However, due to the small sample, there is a possibility that some observable characteristics might "by chance" still be imbalanced between the groups.

As such, when designing the experiment, key demographic characteristics that could influence the prediction of outcomes—such as age, gender, education, etc. —were included in the questions asked to participants, to enable post-experimental balance assessment. Indeed, certain factors like age and gender may act as confounders, potentially shaping how respondents perceive the credibility of ESG disclosures. Notably, research suggests that younger investors, particularly those who are female and have higher household incomes, are more inclined to habe sustainable investment behaviour (Gutsche et al., 2023). This trend is further observed with the growing influence of Millennials in sustainable finance, driven by the anticipated "great wealth transfer "[4], highlighting shifting

[2] It is important to note that, due to the random assignment of treatment, the degree of collinearity induced by the interaction between the expertise and treatment variables should be limited. In our case, this approach can thus be considered conservative. Additionally, we use the term 'collinearity' loosely here, as we are exclusively dealing with binary independent variables. These variables do not correspond to linear measures within a Euclidean space and therefore cannot be collinear in the strictest sense. However, they can still exhibit a form of dependency or association that may contribute to an increase in the variance inflation factor.

[3] Senn (1994) argues that, since randomization ensures balance across treatment and control groups on average, testing for baseline balance only serves to question the integrity of the randomization device itself, not the balance of the study groups.

[4] The "Great Wealth Transfer" refers to the significant money transfer between generations, primarily in the US, as the Baby Boomer generation passes down their assets to their descendants. Cerulli Associates, 2022 predicts that Baby Boomers will transfer $84.4 trillion in assets by 2045, with $72.6 trillion

preferences toward sustainability and social responsibility (Luu and Rubio, 2023). This demographic tends to reflect a value-driven investment philosophy, which may lead to differing perceptions of ESG credibility.

To confirm that our randomization was effective and that the treatment and control groups are balanced, we therefore conduct a balance assessment comparing the means of key demographic characteristics between the two groups (e.g., treatment and control), without to relying on hypothesis testing.

When assessing covariate balance, the literature offers different measures to ensure comparability between groups. One common method is the raw difference in means, which provides a directly interpretable measure of covariate balance for binary variables. However, the literature does not offer specific guidelines (i.e. thresholds) for determining whether these differences indicate meaningful imbalance, and this method is not appropriate when dealing with variables on different scales. The standardized mean difference (SMD) addresses that limitation, accounting for the scale of the variables by standardizing the difference in means, making it more appropriate for comparing continuous variables across different scales.

A further refinement of this approach is the use of the absolute standardized mean difference (ASMD). By focusing on the absolute value of the SMD, the ASMD ensures that the magnitude of the difference, regardless of its direction, is considered. While there is no universally accepted threshold for the standardized difference to indicate the presence of meaningful imbalance (Austin, 2009), guidelines suggest that 0.1 or 0.25 represent reasonable cutoffs for acceptable ASMDs (Stuart et al., 2013). Despite the advantages of using ASMD, ambiguities remain, particularly with binary covariates. The first ambiguity concerns the extent to which covariates with an ASMD around 0.25 can be considered reasonably balanced, since some literature recommends a threshold of 0.1. The second ambiguity involves deciding which measure of (im)balance should be adopted for binary variables, given that general guidelines are available for ASMD but not for raw differences

being directly passed on to their successors.

in means, but at the price of interpretability. Moreover, variables with same imbalances may exhibit widely different SMDs.

To circumvent this problem, our strategy is twofold. First, we present both unstandardized (i.e., simple differences in proportions) and standardized mean differences in the covariates between the treatment and control groups. While the first is easily interpretable, thresholds suggested by the treatment effect literature can be directly applied to the second. Second, we rely on an additional measure that is easily interpretable and for which relatively commonly accepted thresholds exist. This measure is phi (φ), a measure of association between binary variables that Austin, 2009 has shown to be directly related to the ASMD for dichotomous variables. Under the assumption of equal numbers of treated and untreated units, which holds in our case due to random assignment, Austin, 2009 shows that φ can be expressed as a univariate function of the (A)SMD:

$$\varphi_{A2009} = \sqrt{\tfrac{d^2}{40000+d^2}} \# (4)$$

where d is the SMD expressed as a percentage. Since φ gives the association between the covariate and the treatment, and since relevant thresholds have been determined for this measure, we use it to get a rough but easily understandable idea of the severity of the potential bias induced by the difference in mean between treatment and control groups, and the severity of the covariate-specific treatment-control imbalance.

The results of these balance checks serve two main purposes in the context of this study. First, they help assess the extent to which our relatively small sample size could lead to in-sample imbalances, which could “by chance” introduce bias into our estimates despite the random assignment. Second, we face a trade-off between bias and variance reduction: introducing more regressors can help reduce bias but also inflates variance, which reduces statistical power. Therefore, these checks help determine the order in which controls are (if necessary) introduced, or are used in robustness checks to balance this tradeoff effectively.

3.2.3 Adjusting for Heteroskedasticity

While the treatment in our study was randomly assigned, the sampling process itself was not fully randomized. Participants were recruited through clusters—such as students within the same class or professionals connected through LinkedIn networks—introducing the potential for heteroskedasticity and correlated errors within these groups. Ignoring such clustered structures can lead to underestimating standard errors, which could potentially lead to incorrect conclusions about the statistical significance of our results. To formally test for heteroskedasticity, we conduct the Breusch-Pagan and Cook-Weisberg tests (Appendix C). Both confirm the assumption of homoskedasticity is violated, and as such there may be correlated errors among respondents.

One challenge in our study lies in the uncertainty surrounding the exact clustering structure. For example, LinkedIn networks or student group affiliations may create complex, overlapping clusters, which are difficult to capture accurately. Consequently, we cannot define all possible clusters with precision, making it necessary to adopt a flexible approach that acknowledges this uncertainty.

We acknowledge this inherent uncertainty by considering several dimensions and segments of their support when adjusting standard errors for clusters. More precisely, we estimate not one, but several standard errors and P values for each of our coefficient estimates. For each model (each row in the tables), we thus obtain 4 values for both SE and p-value. We finally report the minimum and maximum of these 4 values.

4 Data

4.1 Descriptive Statistics

In total, we collected 221 responses. The characteristics of the participants are detailed in **??** description and Appendix D – Descriptive statistics. The PERCRED score, which serves as the outcome variable of this study, is summarized in Figure 1, while Table 1 provides some descriptive statistics of the PERCRED score and its four constituent factors[5].

[Insert Figure 4.1 about here]

[Insert Table 4.1 about here]

The overall PERCRED score appears to follow a normal distribution, indicating that the respondents' perceptions of credibility are centred around the mean (approximately 4.7), with a relatively balanced spread of responses across the full range from 1 to 7.

The mean scores across the four dimensions (A, S, T, U) and the overall perceived credibility (PERCRED) are relatively close, ranging from 4.4774 to 5.3493. The highest mean is observed in the Understandability (U) dimension at 5.3493, suggesting that respondents generally found the statements easy to comprehend. The standard deviations, ranging from 1.0841 to 1.3282, indicate moderate variability in responses.

The descriptive statistics for the key variables used in the analysis, including the PERCRED scale and associated Likert-scale variables, are presented in Appendix D. This includes a detailed summary of central tendencies, variability, and frequency distributions for variables such as perceived credibility, familiarity with ESG, and report-reading behaviors.

[Insert Table 4.2 about here]

To assess reliability, we employ both Cronbach's alpha and Guttman's lambda 6. Whichever measure of reliability is used, Table 2 shows that in our sample, the reliability of the PERCRED scale is excellent, ensuring that the conclusions drawn from this scale are robust and dependable.

4.2 Balance assessment

We assess the comparability between the treatment and control groups by following the methodology described in section 3.2.2. Specifically, we use descriptive statistics and

[5]The detailed breakdown is in the Appendix D, where all the items are presented along with related descriptive statistics.

present both unstandardized and standardized mean differences in the covariates between the treatment and control groups, as well as the φ measure of association between each covariate and the binary treatment to evaluate the balance of key demographic characteristics.

[Insert Figure 4.2 about here]

The left panel of Figure 2 indicates that none of the covariates possibly impacting the outcome exhibits an ASMD higher than the 0.25 threshold. There is nonetheless a set of variables whose ASMD is exceeding 0.10, which can be considered as the 'safe-side' cutoff. Among these variables, some are closer to an ASMD of 0.25 than 0.10 (with 0.25 being the highest threshold proposed by the literature) and, particularly given the role these covariates play in operationalizing the concept of expertise, we decide to further investigate the balance using φ. In any case, the average ASMD in our sample is 0.12, indicating an acceptable overall balance.

[Insert Figure 4.3 about here]

Given that Figure 3 indicates that the highest φ value in our sample is 0.11, we can confidently conclude that none of our covariates are meaningfully associated with the randomly assigned treatment.

The results shown in Figures 2 and 3 support our decision to adhere to the baseline model, without adjusting for any covariates, except those we assume to mediate the treatment effect, meaning, the variables that act as proxies of expertise. In the separate-sample approach, we effectively control for expertise by conditioning on variables that act as proxies for it, considering only 'expert' individuals. In the pooled-sample approach, we include these expertise proxies both as main effects and through interaction terms[6].

[6] Additional regression analyses where controls for less balanced covariates are introduced are available in Appendix E. These checks are performed to validate the reliability of the study's conclusions, ensuring that the incremental introduction of controls (based on imbalance levels) does not alter the main findings.

5 Results

5.1 Hypothesis Testing

This section details the results of our analysis. Given that '*familiarity*' and *'expertise'* with reporting is an inherently abstract concept that varies significantly among readers, participants were categorized based on their experience with reporting. By operationalizing '*familiarity*' and *'expertise"*into specific, measurable categories, this approach enables an examination of how varying levels of familiarity impact the perceived credibility of precise time-horizon disclosures in ESG reports. The following categories were established:

Familiarity with reporting: This broad category includes any form of experience with either firm or ESG reports, whether through reading, analyzing, or holding a relevant job role.

Familiarity with firm reporting: This subgroup narrows the scope to participants who have either read or analyzed firm reports.

- **Having read a firm report:** Participants who have read at least one firm report.
- **Having analyzed a firm report:** Participants who have gone beyond reading and engaged in detailed analysis of a firm report.

Familiarity with ESG reporting: This subgroup includes participants with experience specifically in ESG reporting, either through reading or analyzing reports or through holding an ESG-related professional role.

- **Having read an ESG report:** Participants who have read at least one ESG report.
- **Having analyzed an ESG report:** Participants who have conducted in-depth analysis of ESG reports.
- **Having an ESG-related profession:** Participants whose professional roles involve ESG-related activities or responsibilities.

[Insert figure 4.4 about here]

Although these subpopulations often overlap, this overlap is intentional. Retaining these overlapping categories enhances robustness by allowing for detailed analysis across multiple dimensions of familiarity, especially in cases of outliers or nuanced interpretations of report reading and analysis.

5.1.1 Separate-sample estimation

We begin by testing Hypothesis 1 (H1), which posits that the effect of precise time-horizon disclosures on the perceived credibility of our population of interest. This hypothesis is tested using a separate-sample approach, as outlined in the Methodology section. Given the challenge of precisely defining the degree of expertise or familiarity that allows the treatment to be effective, we split our sample into various subpopulations, each representing different operationalizations of the concept of expertise as described previously.

[Insert Table 4.3 about here]

As expected, we cannot confidently reject the null hypothesis of no effect for the whole sample (last row), with an estimated ATE of 0.179 and p-values ranging from 0.185 to 0.235. This result is unsurprising as the sample includes individuals with no experience in reporting, who are less likely to notice or appreciate the presence of time-horizon details in ESG reports.

Participants with the strongest treatment effect are those in ESG-oriented occupations, with an ATE of 0.629 and a p-value range of [0.025, 0.043], indicating a statistically significant effect at the 95% confidence level. This finding suggests that professionals with ESG-related expertise are particularly sensitive to the precision of time-horizon information. One potential explanation is that their roles require a deep understanding of how specific details align with broader sustainability goals and regulatory frameworks, making precise timelines essential for assessing the credibility and feasibility of a company's sustainability commitments. Additionally, their professional focus on long-term planning and strategic decision-making may heighten their sensitivity to time-horizon information in ESG disclosures.

When examining subpopulations with varying levels of familiarity with reporting, a less pronounced but still statistically significant effect of the treatment is observed. Participants with any familiarity with reporting (either firm or ESG) show a significant ATE of 0.425 with a p-value range of [0.015, 0.027], demonstrating a significant effect at the 95% confidence level and indicating that even a basic level of familiarity with reporting is sufficient for the treatment to take effect. For participants specifically familiar with firm or ESG reporting, the treatment effect follows a consistent pattern: reading reports yields a significant impact, while analyzing reports does not allow for a confident rejection of the null hypothesis. For those who have read firm reports, the ATE is 0.396, indicating statistical significance at the 95% confidence level. Similarly, for participants who have read ESG reports, the ATE is 0.441, significant at the 90% confidence level. This suggests that simply reading firm or ESG reports enhances the ability to appreciate the precision of time-horizon details.

In contrast, for participants who have analyzed firm or ESG reports, the ATEs drop to 0.317 with a p-value range of [0.195, 0.213] and 0.3 with a p-value range of [0.297, 0.321], respectively, indicating that we cannot confidently reject the null hypothesis of no treatment effect based on this data. This represents an interesting puzzle, as the treatment effect is observed for most groups but not for these two. One potential explanation might lie in the smaller size of these subsamples compared to others, where a small n could explain the high p-values. Another possibility is that while reading reports allows participants to focus on the overall clarity of the information presented—making time-horizon details more impactful—analyzing reports may bring the focus towards critically evaluating specific data points, which can diminish the perceived importance of time-horizon precision. Although we cannot reject the null hypothesis for *Has analyzed firm report* and *Has analyzed ESG report*, the positive direction of the point estimates suggests a practical effect aligned with theoretical expectations, albeit too small to achieve statistical significance within these sample sizes.

In summary, the results indicate that precise time-horizon disclosures significantly enhance perceived credibility among those with ESG-related occupations and those who have read firm or ESG reports, with the strongest effect observed in ESG professionals. This suggests that ESG experts, given their field-specific expectations and reliance on detailed reporting, are particularly sensitive to the clarity that time-horizon precision provides. While general familiarity with reporting, particularly through reading, is sufficient for the treatment to take effect, more in-depth analytical engagement does not necessarily enhance this sensitivity.

5.1.2 Pooled-sample estimation differentiated impact

In this section, we focus on the main effect of the treatment and its interaction with the relevant subpopulations. This analysis incorporates interaction terms to assess the extent to which the effect significantly differs according to the population considered, particularly between experts and non-experts. The total effect for each "expert" group can be determined by summing the associated main effect and the corresponding interaction effect, providing the same result as the separate sample analysis.

[Insert Table 4.4 about here]

The interaction between the treatment and 'Familiarity with reporting' ('Fam. Report' column of Table 4) is significant ($\beta = 0.98$), suggesting that stakeholders familiar with any type of reporting respond more positively than non-experts to precise disclosures, appreciating the enhanced detail due to their broader exposure to report analysis. Interaction effects such as between the treatment and having read a firm report ('Read firm report', $\beta = 0.71$) and an ESG report ('Read ESG report', $\beta = 0.51$) indicate that experience with firm reports particularly enhances sensitivity to precision, with reading having a stronger influence than analyzing. These insights suggest that familiarity with the format and content of firm reports allows stakeholders to better appreciate the implications of detailed, precise information in disclosures.

The interactions involving ESG-specific experiences, like between treatment and 'ESG-related occupation' ($\beta = 0.65$), highlight that professionals whose roles concentrate on ESG issues are particularly attuned to the nuances of precise disclosures, likely due to their focused expertise and the practical implications of such information in their work.

For some interactions however, such as between the treatment and 'Analyzed firm report', as well as with 'Analyzed ESG report', we cannot confidently claim a statistically significant difference in treatment effect, which might suggest that certain types of detailed engagement with ESG content do not necessarily translate to a higher appreciation of precision in disclosures. This could be due to a variety of factors, including the possibility that stakeholders who analyze ESG reports may already expect a high level of detail and precision, thus diminishing the incremental impact of even more precise information.

Overall, the findings support Hypothesis 2 by demonstrating that the effectiveness of precise time-horizon disclosures is contingent on the participant's reporting experience, with significant effects observed among those with general report familiarity and ESG-related professional expertise. However, while not significant, we do unexpectedly observe a negative impact of the treatment on non-experts for some of the moderator variables, which is non expected and partially contradicting H2.

5.2 Participants feedback and observations

At the conclusion of the survey, participants were given the opportunity to share their impressions. Notably, several participants assigned to the vague condition highlighted the absence of precise time horizons in the ESG disclosures. While these observations provide valuable insights, it is important to note that this feedback is more anecdotal in nature. Although the participants' comments align with our hypotheses, they do not represent a comprehensive or systematic analysis. Instead, they offer supplementary qualitative insights that support the broader quantitative findings discussed earlier. For instance, one participant succinctly commented:

> "No deadline for the planned changes."

This comment directly points out the lack of specificity in the timeline, reflecting a common concern among respondents in this group.

Another participant provided a more detailed critique, noting that while the text was "clear" in its objectives, it was vague in the details of implementation. The participant elaborated:

> " In summary, what is "clear":
>
> - The text is understandable regarding the objectives, although typical ("we reduce emissions, we modify our governance"),
> - The text provides two practical "examples" for the environment (oven and hydrogen). That said, the text does not go into detail... (When? What will it replace and how? etc.). [...]
>
> Regarding the trustworthiness of ABC:
>
> - Based on this first reading, it seems like nothing has been put in place in recent years... More trust would be granted if one of the points was to continue the effort already underway (if that were the case...).
> - Also, I don't remember the status of the oven and hydrogen (just a promise?), but more trust would be granted if the "committed status" of these objectives were provided [...]"

Both of these respondents identified themselves as professionals with ESG-oriented backgrounds and previous experience in reading firm reports. While these comments align with the research hypothesis regarding the importance of precision in ESG reporting, it is important to note that this feedback represents anecdotal evidence. Nevertheless, the insights highlight the expectations of professionals who are well-versed in the field, suggesting that they are particularly attuned to the details (or lack thereof) in such reports, reinforcing the need for clarity and specificity to enhance credibility.

6 Concluding remarks

This study explored the impact of time-horizon precision on the perceived credibility of ESG disclosures. Given the research question, an experimental design was particularly relevant to determine the causal impact of varying levels of time-horizon precision on the perceived credibility of ESG disclosures. This approach allowed for the random assignment of treatment conditions, mitigating concerns related to endogeneity and selection bias, and ensuring that any observed effects can be attributed to the manipulation rather than to confounding factors.

The study revealed that the treatment effect is observable in populations with some degrees of familiarity with reporting, whether their connection is direct or indirect. This aligns with Hypothesis H1, which proposed that precise time horizons would enhance the perceived credibility of ESG reports, particularly among stakeholders with relevant experience or familiarity with firm reporting or ESG content. The analysis confirmed that the impact of the treatment is indeed significant within these subpopulations, supporting the notion that expertise and familiarity amplify the effect of time-horizon precision on perceived credibility.

However, when considering the entire sample, the effect of time-horizon precision on perceived credibility could not be confidently distinguished from zero. This was likely due to the inclusion of participants who had no relevant experience with ESG reporting. To further explore this, interaction terms were used in the analysis to examine how different subgroups—specifically experts versus non-experts—responded to time-horizon precision. By incorporating these interaction terms, the study was able to quantify the difference in impact between these groups. The results showed that experts, who were more familiar with reporting practices, had a stronger positive response to precise time-horizon disclosures compared to non-experts. This finding aligned with Hypothesis H2, which suggested that the effectiveness of precise time horizons would vary significantly between experienced and inexperienced individuals.

The observed significant impact among those familiar with firm reports can be intuitively attributed to their prior experience, which provides them with a benchmark for comparison. This familiarity likely enhances their ability to recognize and appreciate the detailed disclosures, leading to a more pronounced positive response. This differential effect underscores the importance of considering audience segmentation in ESG communication strategies. Specifically, stakeholders with prior exposure to firm reports or ESG content are better equipped to discern and value the precision in reporting, which in turn enhances their trust and confidence in the information presented. By showing that there's a direct impact between how precise these disclosures are and how they're perceived, we're emphasizing the need for organizations to adopt clear and transparent reporting practices. Providing clear timelines not only helps build trust and transparency but also strengthens a company's credibility in the eyes of stakeholders. In the long run, this is key for responsible business conduct and driving sustainable development.

7 Figures and Tables

Figure 2.1: PERCRED: histogram and kernel density

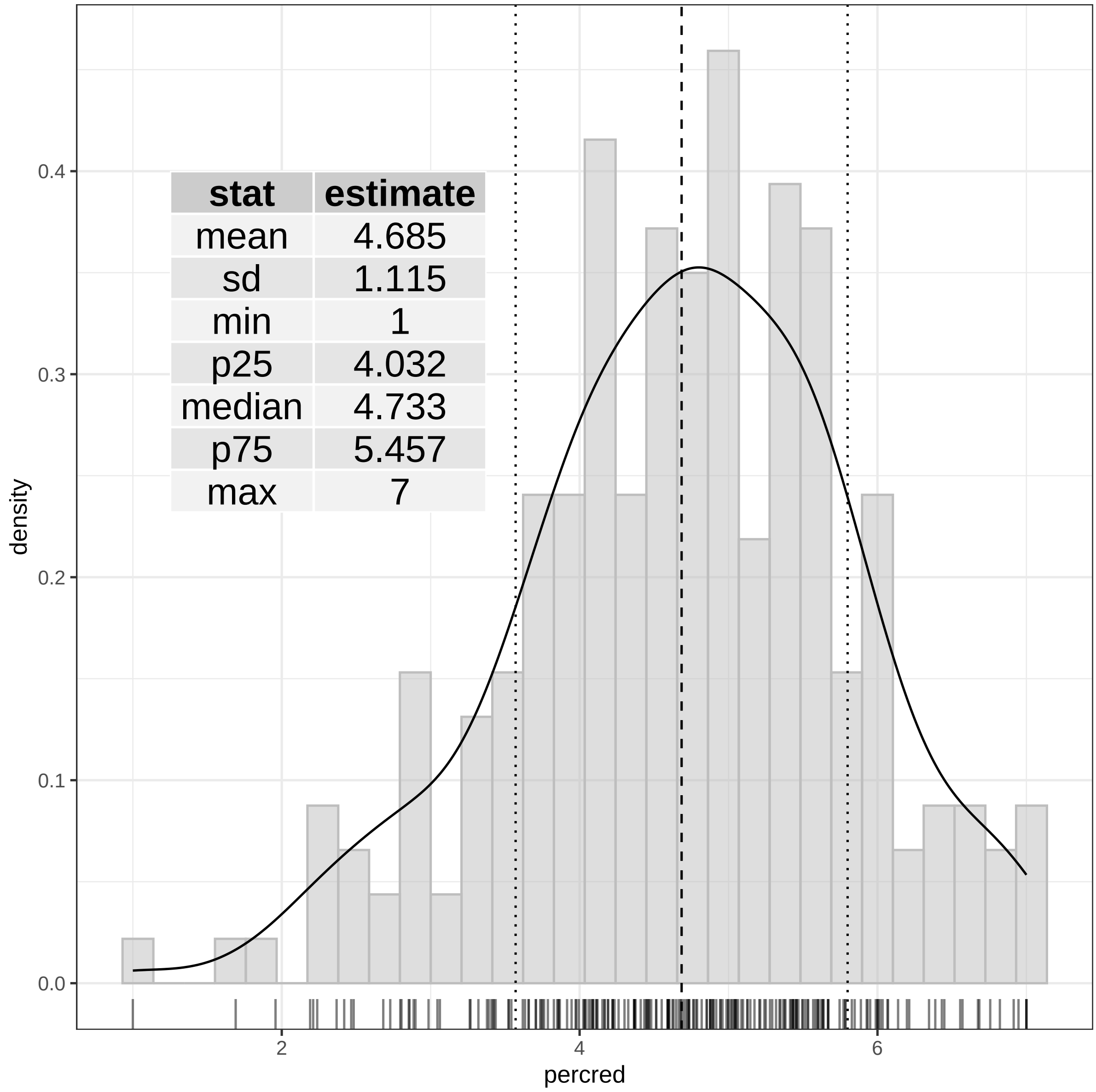


Figure 1 is a density histogram of 'percred' values with summary statistics (mean, standard deviation, min, max, and quartiles). The central dashed line indicates the mean value (4.68), with additional dashed lines showing one standard deviation (±1.11) from the mean on either side. This representation maintains the actual values, enhancing interpretability by illustrating the data's natural spread without standardization. The overlayed density curve and standard deviation markers illustrate the data's spread and central tendency.

Figure 2.2: Mean difference in covariates: raw vs standardized

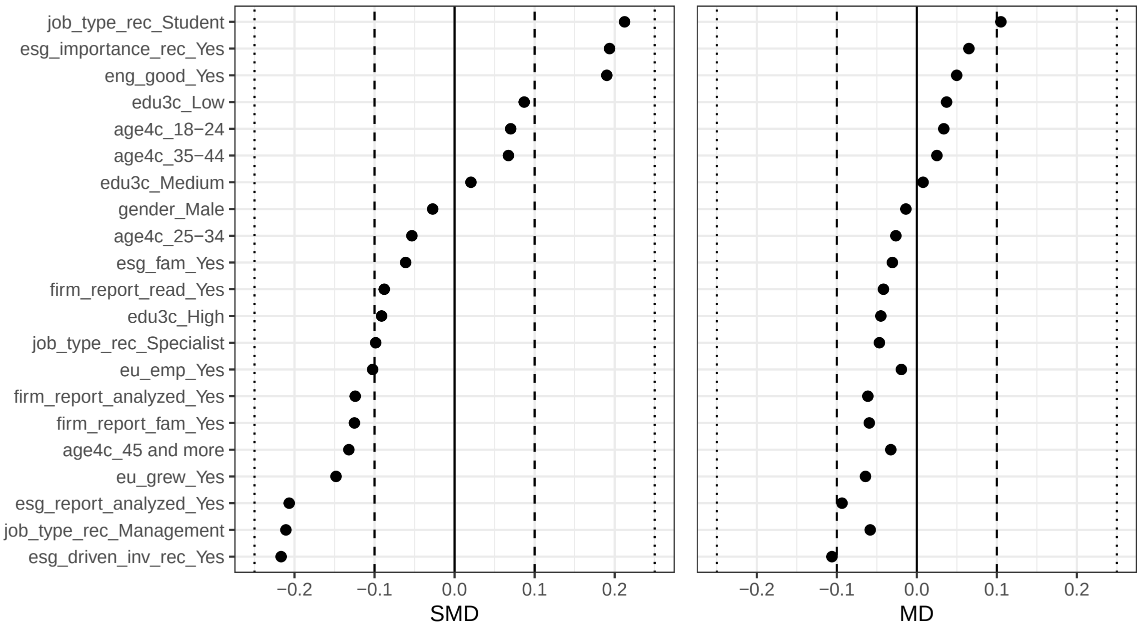


Figure 2 displays both the standardized (left panel) and unstandardized (right panel) mean difference in binary covariates between the treatment and the control group, with covariates arranged in descending order of standardized mean difference (SMD). A comparison of the two panels reveals that standardization slightly changes the order of covariates in the severity of imbalance (see e.g. *job_type_rec_Management* and *eu_emp_Yes*).

Figure 2.3: Assessing balance based on the association between covariates and the binary treatment

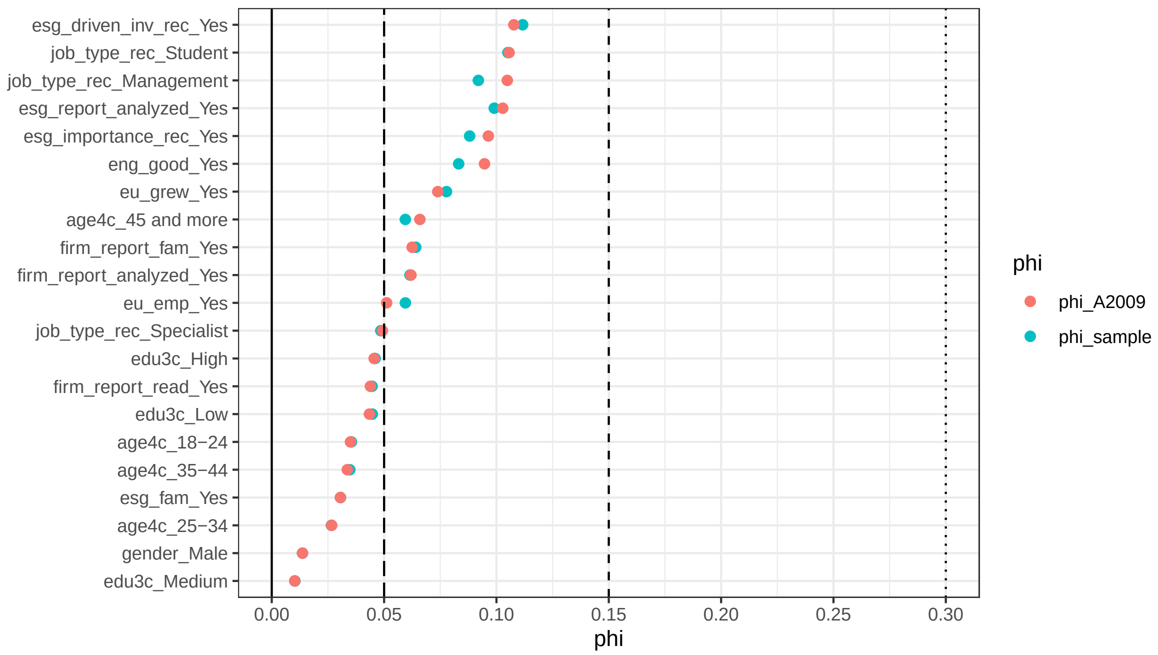


Figure 3 presents the φ measure of association between each covariate and the binary treatment. This figure includes both the value of phi obtained from the SMD, as outlined in Equation 4, and the value of phi directly obtained from the sample according to its standard definition, $\varphi = \sqrt{\frac{\chi_u^2}{N}}$, where χ_u^2 is the uncorrected chi-squared statistic and N is the total sample size. The thresholds represented by the vertical lines are notably derived from the work of Fleiss *et al.* (2003) and Austin (2009): $\varphi \leq 0.3$ implies "no more than trivial association", $\varphi \leq 0.15$ indicates a negligible association, and $\varphi \leq 0.05$ can be considered as indicative of "essentially no correlation" between the treatment group and the binary covariate.

Figure 2.4: Subpopulations categories representing different expertise operationalizations

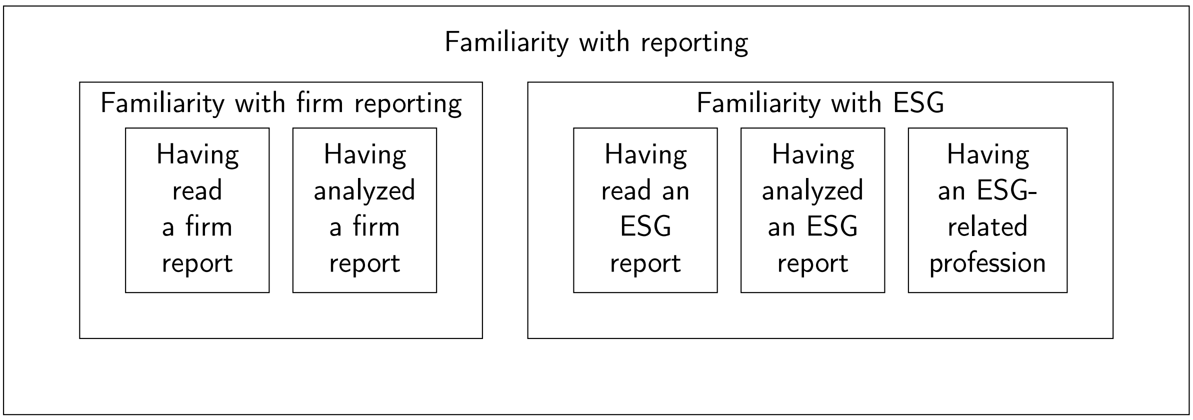


Figure 4 presents the categories of familiarity and expertise used to assess the perceived credibility of time-horizon disclosures in ESG reports. It divides familiarity into two main segments: familiarity with firm reporting and familiarity with ESG reporting. Each segment further breaks down into specific levels of engagement: reading a report, analyzing a report, and for ESG, holding a related professional role. This classification aids in evaluating how different levels of report familiarity affect perceptions among study participants.

Table 2.1: PERCRED and factors: descriptive statistics

Factor/Scale	Label	Mean	SD	Min	P25	Median	P75	Max
A	Appropriateness	4.75	1.21	1	4.00	4.90	4.00	7
S	Sincerity	4.54	1.33	1	3.68	4.66	3.68	7
T	Truth	4.48	1.30	1	3.82	4.62	3.82	7
U	Understandability	5.35	1.08	1	4.74	5.52	4.74	7
PERCRED	Perceived Credibility	4.68	1.11	1	4.03	4.73	4.03	7

Table 1 summarizes the descriptive statistics for each factor, including mean, standard deviation (SD), minimum (Min), and maximum (Max).

Table 2.2: PERCRED and factors: reliability assessment.

Factor/Scale	Alpha (Raw)	Alpha (Std)	Lambda 6
A	0.81	0.83	0.76
S	0.89	0.89	0.85
T	0.92	0.93	0.93
U	0.89	0.90	0.89
PERCRED	0.94	0.94	0.96

Table 2 presents the reliability measures for each factor, including raw alpha (based on covariances), standardized alpha (based on correlations), and Lambda 6.

Table 2.3: Average Treatment effect (ATE) and type and degree of expertise

subpopulation	n	ATE	SE (range)	p-value (range)
ESG-related occupation (Expert)	65	0.629	[0.275, 0.304]	[0.025, 0.043]
Has read an ESG report	107	0.441	[0.233, 0.244]	[0.061, 0.073]
Familiarity with ESG	116	0.432	[0.222, 0.229]	[0.054, 0.062]
Familiarity with reporting	164	0.425	[0.173, 0.191]	[0.015, 0.027]
Familiarity with firm reporting	152	0.400	[0.178, 0.192]	[0.026, 0.039]
Has read a firm report	150	0.396	[0.181, 0.194]	[0.030, 0.043]
Has analyzed a firm report	99	0.317	[0.243, 0.253]	[0.195, 0.213]
Has analyzed an ESG report	74	0.300	[0.286, 0.300]	[0.297, 0.321]
(all)	221	0.179	[0.135, 0.151]	[0.185, 0.235]

Table 3 presents the estimates of the Average Treatment Effect (ATE) across these different subpopulations. The table is organized to show the treatment effect for each subgroup, alongside the associated standard error (SE) and p-value ranges Results are presented in decreasing order of strength of ATE.

Table 2.4: Results of the pooled-sample estimation with interactions terms

moderator	metric	estimate	SE (range)	p-value (range)	n(m)	n(¬m)
ESG-related occupation (Expert)	diff TE	0.650	[0.308, 0.349]	[0.036, 0.064]	65	156
	TE(¬mod)	-0.021	[0.143, 0.170]	[0.884, 0.902]	65	156
Has read an ESG report	diff TE	0.511	[0.291, 0.333]	[0.081, 0.126]	107	114
	TE(¬mod)	-0.070	[0.170, 0.182]	[0.681, 0.701]	107	114
Familiarity with ESG	diff TE	0.549	[0.292, 0.311]	[0.061, 0.079]	116	105
	TE(¬mod)	-0.117	[0.173, 0.190]	[0.500, 0.540]	116	105
Familiarity with reporting	diff TE	0.985	[0.262, 0.312]	[0.000, 0.002]	164	57
	TE(¬mod)	-0.560	[0.199, 0.213]	[0.005, 0.009]	164	57
Familiarity with firm reporting	diff TE	0.763	[0.254, 0.281]	[0.003, 0.007]	152	69
	TE(¬mod)	-0.363	[0.176, 0.205]	[0.040, 0.079]	152	69
Has read a firm report	diff TE	0.711	[0.254, 0.285]	[0.006, 0.013]	150	71
	TE(¬mod)	-0.314	[0.170, 0.209]	[0.066, 0.133]	150	71
Has analyzed a firm report	diff TE	0.299	[0.281, 0.304]	[0.288, 0.327]	99	122
	TE(¬mod)	0.019	[0.163, 0.177]	[0.908, 0.916]	99	122
Has analyzed an ESG report	diff TE	0.239	[0.327, 0.342]	[0.465, 0.485]	74	147
	TE(¬mod)	0.060	[0.147, 0.165]	[0.683, 0.714]	74	147
(none)	ATE	0.179	[0.135, 0.151]	[0.185, 0.235]	221	0

Table 4 displays the differential treatment effects (diff TE) for various moderators representing different levels of expertise and familiarity with ESG or firm reporting. The table lists the estimated effect size (estimate), standard error range (SE), and p-value range for each subgroup, showing how the treatment impact varies across different expertise categories. Results are presented in decreasing order of the strength of the differential treatment effect, highlighting where the treatment effect significantly differs compared to the baseline group.

8 Appendices

	Total		Non-treated		Treated	
	Freq	%	Freq	%	Freq	%
Gender						
Female	111	50.23	56	49.56	55	50.93
Male	110	49.77	57	50.44	53	49.07
Language Proficiency						
Limited proficiency	22	9.95	14	12.39	8	7.41
Native speaker (English is my first language)	10	4.52	5	4.42	5	4.63
Proficient	189	85.52	94	83.19	95	87.96
Country of orgigin						
Europe	169	76.47	89	78.76	80	74.07
Non-Europe	52	23.53	24	21.24	28	25.93
Country of employment						
Europe	208	94.12	106	93.81	102	94.44
Non-Europe	13	5.88	7	6.19	6	5.56
Job Position						
Executive / Senior Management (e.g., CEO, CFO, CTO)	5	2.26	3	2.65	2	1.85
Middle Management (e.g., Department Head, Senior Manager)	20	9.05	13	11.50	7	6.48
Professional / Specialist (e.g., Engineer, Analyst, Lawyer)	81	36.65	44	38.94	37	34.26
Student	115	52.04	53	46.90	62	57.41
Specialization						
ESG-Oriented	65	29.41	37	32.74	28	25.93
Generalist	156	70.59	76	67.26	80	74.07
Firm Report Read						
No	71	32.13	34	30.09	37	34.26
Yes	150	67.87	79	69.91	71	65.74
Firm Report Analyzed						
No	122	55.20	59	52.21	63	58.33
Yes	99	44.80	54	47.79	45	41.67
ESG Report Read						
No	114	51.58	58	51.33	56	51.85
Yes	107	48.42	55	48.67	52	48.15
ESG Report Analyzed						
No	147	66.52	70	61.95	77	71.30
Yes	74	33.48	43	38.05	31	28.70

Table 2.5: Sample description: categorical variables

Variable	**Mean**	**SD**	**Median**	**Min**	**Max**
Age	29.82	9.18	27	18	70
Importance of Sustainability	4.28	0.90	5	1	5
Consider ESG Investment	3.79	1.00	4	1	5

Table 2.6: Sample description: numerical variables

Appendix B - Questionnaire

LimeSurvey

Section C: Section C

Please rate each of the following statements according to your agreement level, 1 being that you completely disagree with the statement and 7 being that you completely agree with the statement.

C1. I understand the text.

	1Completely disagree	2	3	4	5	6	7Completely agree
I understand the text.	☐	☐	☐	☐	☐	☐	☐

C2. The text is clearly written.

	1Completely disagree	2	3	4	5	6	7Completely agree
The text is clearly written.	☐	☐	☐	☐	☐	☐	☐

C3. The text is written in an understandable way.

	1Completely disagree	2	3	4	5	6	7Completely agree
The text is written in an understandable way.	☐	☐	☐	☐	☐	☐	☐

C4. I understand the meaning of the text.

	1Completely disagree	2	3	4	5	6	7Completely agree
I understand the meaning of the text.	☐	☐	☐	☐	☐	☐	☐

C5. The text is easy to read.

	1Completely disagree	2	3	4	5	6	7Completely agree
The text is easy to read.	☐	☐	☐	☐	☐	☐	☐

C6. I think that the statements in the text are accurate;

	1Completely disagree	2	3	4	5	6	7Completely agree
I think that the statements in the text are accurate	☐	☐	☐	☐	☐	☐	☐

C7. I think that the claims made in the text are correct.

	1Completely disagree	2	3	4	5	6	7Completely agree
I think that the claims made in the text are correct.	☐	☐	☐	☐	☐	☐	☐

C8. I am confident that the statements are true.

	1Completely disagree	2	3	4	5	6	7Completely agree
I am confident that the statements are true.	☐	☐	☐	☐	☐	☐	☐

C9. I think that the text uses the best evidence at hand

	1Completely disagree	2	3	4	5	6	7Completely agree
I think that the text uses the best evidence at hand	☐	☐	☐	☐	☐	☐	☐

LimeSurvey

C10. The arguments are justified by the facts in the text.

	1Completely disagree	2	3	4	5	6	7Completely agree
The arguments are justified by the facts in the text.	☐	☐	☐	☐	☐	☐	☐

C11. The text reflects the genuine intentions of the company.

	1Completely disagree	2	3	4	5	6	7Completely agree
The text reflects the genuine intentions of the company.	☐	☐	☐	☐	☐	☐	☐

C12. The text is not misleading.

	1Completely disagree	2	3	4	5	6	7Completely agree
The text is not misleading.	☐	☐	☐	☐	☐	☐	☐

C13. I think that the company's intentions correspond with the text.

	1Completely disagree	2	3	4	5	6	7Completely agree
I think that the company's intentions correspond with the text.	☐	☐	☐	☐	☐	☐	☐

C14. The report fits to the context of the industry and its social and environmental challenges.

	1Completely disagree	2	3	4	5	6	7Completely agree
The report fits to the context of the industry and its social and environmental challenges.	☐	☐	☐	☐	☐	☐	☐

C15. As a reader of this ESG report, I feel that the text addresses ESG issues well

	1Completely disagree	2	3	4	5	6	7Completely agree
As a reader of this ESG report, I feel that the text addresses ESG issues well	☐	☐	☐	☐	☐	☐	☐

C16. I think the text rightfully represents the company.

	1Completely disagree	2	3	4	5	6	7Completely agree
I think the text rightfully represents the company.	☐	☐	☐	☐	☐	☐	☐

Section D: Section D

Thank you for reflecting on the text. Please now take a few moments to answer the following questions.

D1. Which description best fits the focus of your current job position?

- Generalist ☐
- ESG-Oriented ☐

D2. Have you already read a firm's report?

- Yes ☐
- No ☐

LimeSurvey

D3. **Have you already analyzed a firm's report?**

Yes ☐

No ☐

D4. **Have you already read an ESG report ?**

Yes ☐

No ☐

D5. **Have you already analyzed an ESG report ?**

Yes ☐

No ☐

D6. **On a scale of 1 to 5, how important do you believe environmental sustainability is for companies?**

1 being Not Important at All and 5 being Extremely Important

1 ☐

2 ☐

3 ☐

4 ☐

5 ☐

D7. **On a scale of 1 to 5, how likely are you to consider ESG factors when making investment decisions?**

1 being Not likely at All and 5 being Extremely likely

1 ☐

2 ☐

3 ☐

4 ☐

5 ☐

Thank you for participating in our study! Your feedback is valuable to us.

Appendix C – Tests for heteroskedasticity

Table 2.7: Breusch-Pagan test for heteroskedasticity

Moderator	Approach	Statistic	P-value
ESG-related occupation (Expert)	sepsam	1.957	0.162
	poolsam	5.580	0.134
Has read an ESG report	sepsam	2.960	0.085
	poolsam	9.685	0.021
Familiarity with ESG	sepsam	2.587	0.108
	poolsam	8.997	0.029
Familiarity with reporting	sepsam	3.239	0.072
	poolsam	13.167	0.004
Familiarity with firm reporting	sepsam	3.672	0.055
	poolsam	13.320	0.004
Has read a firm report	sepsam	3.645	0.056
	poolsam	12.254	0.007
Has analyzed a firm report	sepsam	2.430	0.119
	poolsam	10.084	0.018
Has analyzed an ESG report	sepsam	1.939	0.164
	poolsam	11.550	0.009
(none)	sepsam	4.532	0.033
	poolsam	4.532	0.033

Note: Null hypothesis of homoskedasticity

Table 2.8: Cook-Weisberg test for heteroskedasticity

Moderator	Approach	Statistic	P-value
ESG-related occupation (Expert)	sepsam	1.240	0.266
	poolsam	6.125	0.106
Has read an ESG report	sepsam	2.537	0.111
	poolsam	10.093	0.018
Familiarity with ESG	sepsam	2.183	0.140
	poolsam	9.305	0.026
Familiarity with reporting	sepsam	2.923	0.087
	poolsam	13.438	0.004
Familiarity with firm reporting	sepsam	3.340	0.068
	poolsam	13.392	0.004
Has read a firm report	sepsam	3.365	0.067
	poolsam	12.404	0.006
Has analyzed a firm report	sepsam	2.027	0.155
	poolsam	10.535	0.015
Has analyzed an ESG report	sepsam	1.616	0.204
	poolsam	12.192	0.007
(none)	sepsam	4.690	0.030
	poolsam	4.690	0.030

Note: Null hypothesis of homoskedasticity

Appendix D - Descriptive statistics

Table 2.9: PERCRED and associated Likert-scale variables

name	label	% NA	mean	sd	min	p25	median	p75	max
percred	Perceived Credibility	0	4.68	1.11	1	4.03	4.73	4.03	7
understood	Report: understood	0	5.64	1.24	1	5.00	6.00	5.00	7
clearly_written	Report: clearly written	0	5.43	1.22	1	5.00	6.00	5.00	7
understandable	Report: written in understandable way	0	5.27	1.33	1	4.00	6.00	4.00	7
understood_meaning	Report: meaning has been understood	0	5.75	1.15	1	5.00	6.00	5.00	7
easy_to_read	Report: easy to read	0	4.72	1.47	1	4.00	5.00	4.00	7
statements_accurate	Report: accurate statements	0	4.82	1.40	1	4.00	5.00	4.00	7
claims_correct	Report: claims are correct	0	4.75	1.33	1	4.00	5.00	4.00	7
statements_true	Report: statements are true	0	4.34	1.45	1	3.00	4.00	3.00	7
best_evidence	Report: uses best evidence	0	4.16	1.60	1	3.00	4.00	3.00	7
factual_arguments	Report: arguments justified by facts	0	4.32	1.58	1	3.00	4.00	3.00	7
genuine_intentions	Report: company's intentions are genuine	0	4.62	1.46	1	4.00	5.00	4.00	7
not_misleading	Report: not misleading	0	4.50	1.50	1	4.00	5.00	4.00	7
intentions_in_text	Report: reflects company's intentions	0	4.50	1.43	1	4.00	5.00	4.00	7
fits_challenges	Report: fits context and challenges	0	5.23	1.22	1	5.00	5.00	5.00	7
esg_addressed	Report: ESG issue are adressed well	0	4.67	1.50	1	4.00	5.00	4.00	7
company_represented	Report: rightfully represents company	0	4.48	1.41	1	4.00	5.00	4.00	7

Note: while 'Report:' variables are Likert-scale variables, they are treated as continuous when linearly combined into factors.

Table 9 presents the Likert-scale variables used to construct the PERCRED scale. The table provides descriptive statistics for each variable, including the percentage of missing values (% NA), mean, standard deviation (sd), minimum (min), 25th percentile (p25), median, 75th percentile (p75), and maximum (max) values. Notably, while these Likert-scale variables are traditionally ordinal, they are treated as continuous for the purpose of constructing the PERCRED scale. This approach allows the extraction of informational content, transforming discrete item responses into a continuous measure of perceived credibility.

Table 2.10: Variables used in models and balance assessments

name	label	n unique	top proportions	levels
age4c	Age (4 categories)	4	25-: 42.1%, 18-: 34.4%, 35-: 15.4%	25-34, 18-24, 35-44, 45 and more
edu3c	Education (3 categories)	3	Hig: 60.6%, Low: 22.2%, Med: 17.2%	High, Low, Medium
eng_good	Good level of English	2	Yes: 90%, No: 10%	Yes, No
esg_driven_inv_rec	ESG-driven investment decisions	2	Yes: 65.6%, No: 34.4%	Yes, No
esg_fam	Familiarity with ESG	2	Yes: 52.5%, No: 47.5%	Yes, No
esg_importance_rec	ESG is important for firms	2	Yes: 83.7%, No: 16.3%	Yes, No
esg_report_analyzed	Already analyzed an ESG report	2	No: 66.5%, Yes: 33.5%	No, Yes
esg_report_read	Already read an ESG report	2	No: 51.6%, Yes: 48.4%	No, Yes
eu_emp	Employed in the EU	2	Yes: 97.3%, No: 2.7%	Yes, No
eu_grew	EU is the region of origin	2	Yes: 78.3%, No: 21.7%	Yes, No
firm_report_analyzed	Already analyzed a firm's report	2	No: 55.2%, Yes: 44.8%	No, Yes
firm_report_fam	Familiarity with firm reporting	2	Yes: 68.8%, No: 31.2%	Yes, No
firm_report_read	Already read a firm's report	2	Yes: 67.9%, No: 32.1%	Yes, No
gender	Gender	2	Fem: 50.2%, Mal: 49.8%	Female, Male
job_esg	ESG-related occupation	2	No: 70.6%, Yes: 29.4%	No, Yes
job_type_rec	Job type/level	3	Stu: 52%, Spe: 36.7%, Man: 11.3%	Student, Specialist, Management
report_fam	Familiarity with reporting (firm/ESG)	2	Yes: 74.2%, No: 25.8%	Yes, No
treated	Received the treatment	2	No: 51.1%, Yes: 48.9%	No, Yes

Table 10 outlines the categorical variables employed in our analytical models and balance assessments. These variables range from multinomial and ordered categories to binary dummies, reflecting adaptations made to accommodate distribution characteristics and enhance representativeness. For instance, the 'age' variable was originally a discrete numeric variable. However, to ensure a more balanced representation across groups, it was recoded into four categories. Similarly, other variables such as education levels, familiarity with ESG factors, and employment status were also categorized to facilitate our analyses. The table lists each variable's name, label, the number of unique categories (n unique), the top proportions of responses, and the distinct levels into which responses are categorized. This recoding strategy not only aids in addressing uneven distributions but also enhances the interpretability and relevance of the models to varied demographic and professional profiles.

Appendix E - Robustness checks

The following tables present the results of the separate and pooled-sample regression analysis, where controls are introduced for the less balanced covariates. These controls are introduced incrementally, in decreasing order of imbalance as measured by ASMD, using stepwise regressions. Only covariates with an ASMD greater than 0.1 are included, as those below this threshold are already considered sufficiently balanced. Including additional controls would inflate standard errors and unnecessarily reduce the power of the tests, particularly given the small sample size.

In terms of robustness, the inclusion of controls does not invalidate the conclusions drawn from the baseline model (except for *treatedYes report readYes*, where the null hypothesis cannot be rejected in the final model— which is expected given the progressively smaller sample size as more controls are added). Overall, the results remain robust, reinforcing the positive effect of time-horizon precision on perceived credibility, particularly among individuals familiar with ESG reporting.

	1	2	3	4	5	6	7	8	9	10	11
(Intercept)	4.35***	4.50***	4.86***	4.99***	5.20***	5.42***	5.48***	5.78***	5.78***	5.59***	6.18***
	(0.17)	(0.29)	(0.30)	(0.30)	(0.40)	(0.73)	(0.77)	(0.74)	(0.76)	(0.74)	(1.05)
treatedYes	0.43°	0.42°	0.37°	0.34	0.37	0.37	0.35	0.42°	0.42°	0.44°	0.42°
	(0.23)	(0.23)	(0.22)	(0.22)	(0.23)	(0.23)	(0.24)	(0.23)	(0.23)	(0.23)	(0.23)
esg_driven_inv_recYes		−0.20	0.00	0.03	0.04	0.03	0.03	−0.01	−0.01	−0.00	0.00
		(0.28)	(0.27)	(0.28)	(0.28)	(0.29)	(0.29)	(0.30)	(0.31)	(0.31)	(0.32)
job_type_recManagement			−0.84*	−0.81*	−0.82*	−0.78*	−0.77*	−0.18	−0.18	−0.09	−0.15
			(0.37)	(0.37)	(0.37)	(0.38)	(0.39)	(0.51)	(0.54)	(0.55)	(0.56)
job_type_recSpecialist			−0.73**	−0.68*	−0.67*	−0.64*	−0.63*	−0.23	−0.23	−0.27	−0.28
			(0.26)	(0.27)	(0.28)	(0.29)	(0.29)	(0.32)	(0.37)	(0.38)	(0.38)
esg_report_analyzedYes				−0.25	−0.26	−0.25	−0.25	−0.35	−0.35	−0.10	−0.10
				(0.23)	(0.24)	(0.24)	(0.24)	(0.24)	(0.25)	(0.31)	(0.33)
esg_importance_recYes					−0.27	−0.28	−0.29	−0.35	−0.35	−0.24	−0.25
					(0.37)	(0.38)	(0.37)	(0.37)	(0.37)	(0.39)	(0.39)
eng_goodYes						−0.23	−0.20	0.02	0.02	0.12	0.08
						(0.63)	(0.62)	(0.56)	(0.59)	(0.53)	(0.53)
eu_grewYes							−0.12	−0.18	−0.18	−0.20	−0.14
							(0.25)	(0.26)	(0.26)	(0.26)	(0.27)
age4c25-34								−0.75*	−0.75*	−0.74*	−0.72*
								(0.32)	(0.32)	(0.32)	(0.33)
age4c35-44								−0.84°	−0.84°	−0.83°	−0.82°
								(0.43)	(0.43)	(0.43)	(0.44)
age4c45 and more								−1.14*	−1.14*	−1.11°	−1.05°
								(0.57)	(0.57)	(0.56)	(0.57)
firm_report_famYes									−0.01	0.18	0.17
									(0.41)	(0.43)	(0.43)
firm_report_analyzedYes										−0.49	−0.51
										(0.32)	(0.33)
eu_empYes											−0.59
											(0.68)
R^2	0.03	0.04	0.12	0.13	0.14	0.14	0.14	0.19	0.19	0.21	0.22
Adj. R^2	0.02	0.02	0.09	0.09	0.09	0.08	0.07	0.11	0.10	0.11	0.11
Num. obs.	116	116	116	116	116	116	116	116	116	116	116

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\circ}p < 0.1$

Table 1: Familiarity with ESG

	1	2	3	4	5	6	7	8
(Intercept)	4.26^{***}	4.44^{***}	4.91^{***}	4.91^{***}	4.98^{***}	4.89^{**}	4.88^{**}	5.11^{**}
	(0.22)	(0.45)	(0.49)	(0.49)	(0.62)	(1.57)	(1.63)	(1.66)
treatedYes	0.30	0.30	0.17	0.17	0.18	0.18	0.18	0.21
	(0.30)	(0.30)	(0.31)	(0.31)	(0.32)	(0.33)	(0.33)	(0.32)
esg_driven_inv_recYes		−0.22	−0.05	−0.05	−0.04	−0.03	−0.04	−0.27
		(0.43)	(0.45)	(0.45)	(0.47)	(0.47)	(0.48)	(0.51)
job_type_recManagement			-1.02°	-1.02°	-1.02°	-1.03°	-1.03°	−0.42
			(0.56)	(0.56)	(0.56)	(0.57)	(0.58)	(0.77)
job_type_recSpecialist			−0.68	−0.68	−0.69	−0.69	−0.70	−0.33
			(0.42)	(0.42)	(0.42)	(0.43)	(0.44)	(0.48)
esg_importance_recYes					−0.09	−0.09	−0.08	0.00
					(0.55)	(0.55)	(0.58)	(0.59)
eng_goodYes						0.09	0.08	0.13
						(1.42)	(1.42)	(1.43)
eu_grewYes							0.03	0.11
							(0.35)	(0.38)
age4c25-34								−0.60
								(0.43)
age4c35-44								−0.71
								(0.59)
age4c45 and more								−1.36
								(0.94)
R^2	0.01	0.02	0.09	0.09	0.09	0.09	0.09	0.14
Adj. R^2	−0.00	−0.01	0.04	0.04	0.03	0.01	−0.00	0.01
Num. obs.	74	74	74	74	74	74	74	74

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\circ}p < 0.1$

Table 2: Has analyzed an ESG report

	1	2	3	4	5	6	7	8	9	10	11
(Intercept)	4.27^{***}	4.36^{***}	4.75^{***}	4.84^{***}	5.09^{***}	5.26^{***}	5.28^{***}	5.61^{***}	5.57^{***}	5.30^{***}	6.10^{***}
	(0.18)	(0.31)	(0.34)	(0.35)	(0.42)	(0.82)	(0.85)	(0.80)	(0.83)	(0.81)	(1.07)
treatedYes	0.44°	0.44°	0.38	0.36	0.39	0.40	0.39	0.44°	0.44°	0.46°	0.44°
	(0.24)	(0.24)	(0.24)	(0.24)	(0.25)	(0.26)	(0.26)	(0.25)	(0.25)	(0.25)	(0.25)
esg_driven_inv_recYes		-0.11	0.04	0.06	0.08	0.07	0.07	0.06	0.07	0.08	0.09
		(0.29)	(0.29)	(0.30)	(0.31)	(0.31)	(0.31)	(0.32)	(0.33)	(0.33)	(0.34)
job_type_recManagement			-0.87^{*}	-0.87^{*}	-0.87^{*}	-0.84^{*}	-0.83^{*}	-0.25	-0.26	-0.16	-0.24
			(0.38)	(0.39)	(0.38)	(0.40)	(0.41)	(0.54)	(0.55)	(0.56)	(0.57)
job_type_recSpecialist			-0.66^{*}	-0.64^{*}	-0.62^{*}	-0.60°	-0.59°	-0.20	-0.21	-0.25	-0.26
			(0.29)	(0.30)	(0.31)	(0.32)	(0.33)	(0.36)	(0.38)	(0.39)	(0.39)
esg_report_analyzedYes				-0.15	-0.15	-0.14	-0.14	-0.25	-0.26	0.08	0.10
				(0.25)	(0.25)	(0.26)	(0.26)	(0.25)	(0.26)	(0.30)	(0.32)
esg_importance_recYes					-0.33	-0.35	-0.35	-0.41	-0.42	-0.27	-0.28
					(0.39)	(0.39)	(0.39)	(0.39)	(0.39)	(0.41)	(0.41)
eng_goodYes						-0.18	-0.16	0.03	0.01	0.15	0.10
						(0.74)	(0.74)	(0.66)	(0.73)	(0.63)	(0.63)
eu_grewYes							-0.05	-0.11	-0.11	-0.12	-0.03
							(0.25)	(0.27)	(0.28)	(0.27)	(0.27)
age4c25-34								-0.76^{*}	-0.75^{*}	-0.76^{*}	-0.73^{*}
								(0.34)	(0.35)	(0.36)	(0.37)
age4c35-44								-0.90°	-0.89°	-0.87°	-0.87°
								(0.47)	(0.48)	(0.48)	(0.48)
age4c45 and more								-1.07°	-1.06°	-1.01°	-0.93
								(0.58)	(0.59)	(0.58)	(0.59)
firm_report_famYes									0.07	0.29	0.25
									(0.62)	(0.64)	(0.66)
firm_report_analyzedYes										-0.64^{*}	-0.67^{*}
										(0.31)	(0.32)
eu_empYes											-0.78
											(0.62)
R^2	0.03	0.03	0.11	0.11	0.12	0.12	0.12	0.17	0.17	0.20	0.22
Adj. R^2	0.02	0.02	0.07	0.07	0.07	0.06	0.05	0.08	0.07	0.09	0.10
Num. obs.	107	107	107	107	107	107	107	107	107	107	107

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\circ}p < 0.1$

Table 3: Has read an ESG report

	1	2	3	4	5	6	7	8
(Intercept)	4.29***	4.50***	4.73***	4.85***	4.91***	4.57***	4.61***	4.67***
	(0.18)	(0.29)	(0.31)	(0.32)	(0.36)	(0.65)	(0.68)	(0.72)
treatedYes	0.32	0.30	0.24	0.23	0.24	0.24	0.24	0.21
	(0.25)	(0.25)	(0.25)	(0.25)	(0.26)	(0.26)	(0.27)	(0.28)
esg_driven_inv_recYes		−0.28	−0.13	−0.08	−0.07	−0.04	−0.02	−0.11
		(0.28)	(0.30)	(0.32)	(0.32)	(0.34)	(0.35)	(0.36)
job_type_recManagement			−0.56	−0.55	−0.54	−0.57	−0.57	−0.24
			(0.39)	(0.39)	(0.39)	(0.40)	(0.41)	(0.56)
job_type_recSpecialist			−0.45	−0.41	−0.40	−0.42	−0.42	−0.25
			(0.30)	(0.32)	(0.32)	(0.33)	(0.34)	(0.41)
esg_report_analyzedYes				−0.26	−0.24	−0.27	−0.26	−0.30
				(0.29)	(0.31)	(0.31)	(0.31)	(0.31)
esg_importance_recYes					−0.10	−0.14	−0.16	−0.05
					(0.37)	(0.38)	(0.40)	(0.42)
eng_goodYes						0.41	0.45	0.42
						(0.56)	(0.57)	(0.57)
eu_grewYes							−0.10	−0.03
							(0.30)	(0.31)
age4c25-34								−0.22
								(0.41)
age4c35-44								−0.27
								(0.49)
age4c45 and more								−0.79
								(0.69)
R^2	0.02	0.03	0.06	0.07	0.07	0.07	0.08	0.09
Adj. R^2	0.01	0.01	0.02	0.02	0.01	0.00	−0.01	−0.02
Num. obs.	99	99	99	99	99	99	99	99

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\circ}p < 0.1$

Table 4: Has analyzed a firm report

	1	2	3	4	5	6	7	8	9	10	11
(Intercept)	4.40^{***}	4.48^{***}	4.73^{***}	4.81^{***}	4.84^{***}	5.04^{***}	5.12^{***}	5.24^{***}	5.24^{***}	5.42^{***}	6.37^{***}
	(0.15)	(0.21)	(0.23)	(0.23)	(0.26)	(0.55)	(0.57)	(0.57)	(0.57)	(0.58)	(0.77)
treatedYes	0.40^{*}	0.39^{*}	0.35°	0.33°	0.34°	0.34°	0.33	0.34°	0.34°	0.35°	0.32
	(0.19)	(0.20)	(0.19)	(0.19)	(0.20)	(0.20)	(0.20)	(0.20)	(0.20)	(0.20)	(0.20)
esg_driven_inv_recYes		-0.12	0.06	0.15	0.15	0.14	0.16	0.13	0.13	0.15	0.15
		(0.21)	(0.22)	(0.23)	(0.24)	(0.24)	(0.24)	(0.25)	(0.25)	(0.25)	(0.25)
job_type_recManagement			-0.74^{*}	-0.71^{*}	-0.71^{*}	-0.68^{*}	-0.68^{*}	-0.41	-0.41	-0.36	-0.45
			(0.32)	(0.32)	(0.32)	(0.33)	(0.33)	(0.44)	(0.44)	(0.44)	(0.45)
job_type_recSpecialist			-0.51^{*}	-0.47^{*}	-0.47°	-0.45°	-0.45°	-0.22	-0.22	-0.25	-0.27
			(0.22)	(0.23)	(0.24)	(0.24)	(0.24)	(0.30)	(0.30)	(0.30)	(0.30)
esg_report_analyzedYes				-0.30	-0.30	-0.29	-0.31	-0.33	-0.33	-0.16	-0.14
				(0.21)	(0.21)	(0.22)	(0.22)	(0.22)	(0.22)	(0.25)	(0.25)
esg_importance_recYes					-0.05	-0.04	-0.05	-0.02	-0.02	-0.04	-0.08
					(0.26)	(0.27)	(0.27)	(0.29)	(0.29)	(0.28)	(0.28)
eng_goodYes						-0.23	-0.13	-0.07	-0.07	-0.06	-0.14
						(0.55)	(0.54)	(0.52)	(0.52)	(0.49)	(0.49)
eu_grewYes							-0.22	-0.25	-0.25	-0.29	-0.19
							(0.24)	(0.24)	(0.24)	(0.24)	(0.24)
age4c25-34								-0.39	-0.39	-0.41	-0.36
								(0.30)	(0.30)	(0.29)	(0.30)
age4c35-44								-0.44	-0.44	-0.43	-0.39
								(0.38)	(0.38)	(0.38)	(0.38)
age4c45 and more								-0.35	-0.35	-0.34	-0.25
								(0.52)	(0.52)	(0.50)	(0.51)
firm_report_analyzedYes										-0.37	-0.39°
										(0.23)	(0.23)
eu_empYes											-0.94°
											(0.53)
R^2	0.03	0.03	0.08	0.09	0.09	0.10	0.10	0.11	0.11	0.13	0.15
Adj. R^2	0.02	0.02	0.06	0.06	0.06	0.05	0.05	0.05	0.05	0.05	0.07
Num. obs.	152	152	152	152	152	152	152	152	152	152	152

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\circ}p < 0.1$

Table 5: Familiarity with firm reporting

	1	2	3	4	5	6	7	8
(Intercept)	4.41^{***}	4.50^{***}	4.76^{***}	4.85^{***}	4.82^{***}	4.87^{***}	4.95^{***}	5.05^{***}
	(0.15)	(0.22)	(0.23)	(0.23)	(0.27)	(0.63)	(0.68)	(0.66)
treatedYes	0.40^{*}	0.38°	0.34°	0.32°	0.32	0.32	0.31	0.32
	(0.19)	(0.20)	(0.19)	(0.19)	(0.20)	(0.20)	(0.20)	(0.20)
esg_driven_inv_recYes		−0.12	0.07	0.15	0.15	0.14	0.16	0.14
		(0.21)	(0.22)	(0.24)	(0.24)	(0.24)	(0.25)	(0.26)
job_type_recManagement			-0.77^{*}	-0.75^{*}	-0.75^{*}	-0.74^{*}	-0.73^{*}	−0.50
			(0.32)	(0.32)	(0.33)	(0.33)	(0.34)	(0.45)
job_type_recSpecialist			-0.56^{*}	-0.52^{*}	-0.52^{*}	-0.51^{*}	-0.51^{*}	−0.31
			(0.23)	(0.23)	(0.24)	(0.25)	(0.25)	(0.31)
esg_report_analyzedYes				−0.30	−0.30	−0.30	−0.31	−0.33
				(0.21)	(0.21)	(0.22)	(0.22)	(0.22)
esg_importance_recYes					0.04	0.03	0.02	0.05
					(0.27)	(0.27)	(0.28)	(0.29)
eng_goodYes						−0.05	0.01	0.08
						(0.61)	(0.60)	(0.57)
eu_grewYes							−0.18	−0.21
							(0.24)	(0.25)
age4c25-34								−0.36
								(0.30)
age4c35-44								−0.39
								(0.38)
age4c45 and more								−0.32
								(0.52)
R^2	0.03	0.03	0.09	0.10	0.10	0.10	0.11	0.12
Adj. R^2	0.02	0.02	0.06	0.07	0.06	0.06	0.06	0.05
Num. obs.	150	150	150	150	150	150	150	150

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\circ}p < 0.1$

Table 6: Has read a firm report

	1	2	3	4	5	6	7	8	9	10	11
(Intercept)	4.28***	4.59***	5.28***	5.37***	5.71***	5.39***	5.58***	5.85***	5.42***	5.30***	6.56***
	(0.22)	(0.47)	(0.32)	(0.35)	(0.54)	(1.07)	(1.20)	(1.21)	(1.16)	(1.10)	(1.25)
treatedYes	0.63*	0.58°	0.53*	0.53*	0.63*	0.64*	0.65*	0.73*	0.71*	0.75*	0.67*
	(0.30)	(0.33)	(0.26)	(0.27)	(0.29)	(0.31)	(0.31)	(0.31)	(0.31)	(0.32)	(0.32)
esg_driven_inv_recYes		−0.34	0.12	0.13	0.21	0.25	0.29	0.27	0.29	0.25	0.21
		(0.43)	(0.27)	(0.26)	(0.27)	(0.28)	(0.30)	(0.33)	(0.35)	(0.34)	(0.34)
job_type_recManagement			−1.48***	−1.44**	−1.51**	−1.58***	−1.58**	−0.79	−0.96	−0.96	−1.21
			(0.43)	(0.45)	(0.47)	(0.45)	(0.46)	(0.78)	(0.80)	(0.81)	(0.84)
job_type_recSpecialist			−1.46***	−1.41***	−1.43***	−1.49***	−1.44***	−1.09**	−1.32**	−1.38**	−1.44**
			(0.31)	(0.35)	(0.35)	(0.32)	(0.33)	(0.40)	(0.42)	(0.44)	(0.43)
esg_report_analyzedYes				−0.18	−0.21	−0.20	−0.25	−0.48	−0.80°	−0.44	−0.51
				(0.35)	(0.34)	(0.35)	(0.36)	(0.38)	(0.45)	(0.65)	(0.71)
esg_importance_recYes					−0.44	−0.44	−0.53	−0.55	−0.49	−0.44	−0.47
					(0.54)	(0.54)	(0.55)	(0.55)	(0.55)	(0.53)	(0.52)
eng_goodYes						0.33	0.38	0.48	0.48	0.66	0.60
						(0.93)	(0.92)	(0.87)	(0.90)	(0.81)	(0.82)
eu_grewYes							−0.25	−0.23	−0.13	−0.20	−0.10
							(0.33)	(0.34)	(0.35)	(0.35)	(0.36)
age4c25-34								−0.53	−0.57	−0.61°	−0.62°
								(0.35)	(0.34)	(0.36)	(0.36)
age4c35-44								−0.80	−0.85°	−0.82	−0.84
								(0.48)	(0.51)	(0.51)	(0.51)
age4c45 and more								−1.50°	−1.75°	−1.62°	−1.50
								(0.88)	(0.92)	(0.89)	(0.99)
firm_report_famYes									0.84	1.09	1.18
									(0.65)	(0.69)	(0.71)
firm_report_analyzedYes										−0.65	−0.64
										(0.57)	(0.59)
eu_empYes											−1.20*
											(0.53)
R^2	0.06	0.07	0.33	0.33	0.34	0.34	0.35	0.40	0.41	0.44	0.47
Adj. R^2	0.05	0.04	0.28	0.27	0.27	0.26	0.26	0.27	0.28	0.30	0.32
Num. obs.	65	65	65	65	65	65	65	65	65	65	65

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\circ}p < 0.1$

Table 7: ESG-related occupation (Expert)

	1	2	3	4	5	6	7	8	9	10	11
(Intercept)	4.42***	4.49***	4.76***	4.82***	4.86***	5.04***	5.10***	5.25***	5.31***	5.33***	6.30***
	(0.14)	(0.21)	(0.21)	(0.21)	(0.26)	(0.47)	(0.50)	(0.50)	(0.57)	(0.57)	(0.78)
treatedYes	0.42*	0.42*	0.37*	0.35°	0.36°	0.36°	0.35°	0.36°	0.36°	0.36°	0.34°
	(0.18)	(0.19)	(0.18)	(0.18)	(0.18)	(0.19)	(0.19)	(0.19)	(0.19)	(0.19)	(0.19)
esg_driven_inv_recYes		−0.08	0.07	0.14	0.15	0.13	0.14	0.12	0.11	0.13	0.13
		(0.20)	(0.20)	(0.21)	(0.21)	(0.22)	(0.22)	(0.23)	(0.23)	(0.23)	(0.24)
job_type_recManagement			−0.83**	−0.78**	−0.77*	−0.74*	−0.74*	−0.49	−0.48	−0.43	−0.52
			(0.29)	(0.30)	(0.30)	(0.30)	(0.31)	(0.41)	(0.42)	(0.43)	(0.43)
job_type_recSpecialist			−0.55**	−0.49*	−0.49*	−0.47*	−0.46*	−0.24	−0.22	−0.25	−0.28
			(0.20)	(0.22)	(0.22)	(0.22)	(0.23)	(0.27)	(0.28)	(0.29)	(0.29)
esg_report_analyzedYes				−0.30	−0.29	−0.29	−0.29	−0.33	−0.32	−0.15	−0.13
				(0.21)	(0.21)	(0.21)	(0.21)	(0.22)	(0.22)	(0.25)	(0.25)
esg_importance_recYes					−0.06	−0.05	−0.05	−0.03	−0.03	−0.05	−0.09
					(0.25)	(0.25)	(0.25)	(0.26)	(0.27)	(0.26)	(0.26)
eng_goodYes						−0.20	−0.15	−0.06	−0.05	−0.04	−0.10
						(0.43)	(0.42)	(0.41)	(0.42)	(0.40)	(0.40)
eu_grewYes							−0.15	−0.21	−0.21	−0.25	−0.16
							(0.22)	(0.23)	(0.23)	(0.23)	(0.23)
age4c25-34								−0.44°	−0.45°	−0.47°	−0.41
								(0.27)	(0.27)	(0.27)	(0.27)
age4c35-44								−0.42	−0.42	−0.41	−0.37
								(0.36)	(0.36)	(0.36)	(0.36)
age4c45 and more								−0.43	−0.44	−0.43	−0.33
								(0.49)	(0.50)	(0.48)	(0.49)
firm_report_famYes									−0.08	0.09	0.06
									(0.33)	(0.34)	(0.34)
firm_report_analyzedYes										−0.36	−0.38°
										(0.23)	(0.23)
eu_empYes											−0.95°
											(0.55)
R^2	0.03	0.03	0.10	0.11	0.11	0.12	0.12	0.14	0.14	0.15	0.17
Adj. R^2	0.03	0.02	0.08	0.09	0.08	0.08	0.07	0.07	0.07	0.08	0.09
Num. obs.	164	164	164	164	164	164	164	164	164	164	164

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\circ}p < 0.1$

Table 8: Familiarity with reporting

	1	2	3	4	5	6	7	8	9	10	11
(Intercept)	4.89***	4.94***	5.08***	5.06***	5.20***	5.42***	5.51***	5.59***	5.57***	5.61***	6.57***
	(0.14)	(0.20)	(0.20)	(0.20)	(0.23)	(0.29)	(0.34)	(0.34)	(0.35)	(0.35)	(0.59)
treatedYes	−0.12	−0.13	−0.19	−0.18	−0.17	−0.13	−0.13	−0.16	−0.16	−0.17	−0.16
	(0.19)	(0.20)	(0.20)	(0.20)	(0.20)	(0.20)	(0.20)	(0.21)	(0.21)	(0.22)	(0.21)
esg_famYes	−0.54*	−0.53*	−0.37	−0.17	−0.15	−0.12	−0.13	−0.12	−0.13	−0.17	−0.13
	(0.22)	(0.22)	(0.23)	(0.26)	(0.26)	(0.26)	(0.26)	(0.26)	(0.27)	(0.27)	(0.27)
treatedYes:esg_famYes	0.55°	0.56°	0.57°	0.53°	0.53°	0.49	0.47	0.51°	0.51	0.54°	0.50
	(0.29)	(0.30)	(0.29)	(0.29)	(0.29)	(0.30)	(0.30)	(0.31)	(0.31)	(0.31)	(0.31)
esg_driven_inv_recYes		−0.08	−0.01	0.01	0.04	0.02	0.04	0.03	0.04	0.05	0.05
		(0.17)	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)	(0.18)	(0.18)	(0.18)
job_type_recManagement			−0.71*	−0.69*	−0.69*	−0.65*	−0.63*	−0.55	−0.56	−0.51	−0.59
			(0.27)	(0.27)	(0.27)	(0.28)	(0.28)	(0.35)	(0.36)	(0.37)	(0.37)
job_type_recSpecialist			−0.47**	−0.43*	−0.42*	−0.39*	−0.38*	−0.27	−0.29	−0.31	−0.33
			(0.18)	(0.18)	(0.18)	(0.18)	(0.19)	(0.22)	(0.22)	(0.22)	(0.22)
esg_report_analyzedYes				−0.30	−0.31	−0.30	−0.30	−0.33	−0.35	−0.16	−0.15
				(0.23)	(0.23)	(0.23)	(0.23)	(0.23)	(0.23)	(0.26)	(0.27)
esg_importance_recYes					−0.21	−0.20	−0.20	−0.19	−0.19	−0.20	−0.24
					(0.19)	(0.19)	(0.19)	(0.19)	(0.20)	(0.20)	(0.19)
eng_goodYes						−0.30	−0.26	−0.20	−0.21	−0.20	−0.23
						(0.25)	(0.25)	(0.26)	(0.26)	(0.25)	(0.25)
eu_grewYes							−0.14	−0.19	−0.18	−0.21	−0.14
							(0.20)	(0.20)	(0.20)	(0.20)	(0.20)
age4c25-34								−0.27	−0.26	−0.28	−0.24
								(0.19)	(0.19)	(0.19)	(0.20)
age4c35-44								−0.12	−0.12	−0.11	−0.09
								(0.29)	(0.29)	(0.29)	(0.29)
age4c45 and more								−0.27	−0.27	−0.27	−0.19
								(0.41)	(0.41)	(0.40)	(0.40)
firm_report_famYes									0.06	0.22	0.20
									(0.18)	(0.21)	(0.21)
firm_report_analyzedYes										−0.38°	−0.40°
										(0.23)	(0.23)
eu_empYes											−0.99*
											(0.49)
R^2	0.04	0.04	0.09	0.10	0.10	0.11	0.11	0.12	0.12	0.13	0.15
Adj. R^2	0.02	0.02	0.06	0.07	0.07	0.07	0.07	0.06	0.06	0.07	0.09
Num. obs.	221	221	221	221	221	221	221	221	221	221	221

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\circ}p < 0.1$

Table 1: Familiarity with ESG

	1	2	3	4	5	6	7	8	9	10	11
(Intercept)	4.80***	4.83***	4.96***	4.96***	5.10***	5.36***	5.48***	5.55***	5.52***	5.55***	6.53***
	(0.13)	(0.19)	(0.18)	(0.18)	(0.22)	(0.29)	(0.34)	(0.35)	(0.36)	(0.36)	(0.57)
treatedYes	0.06	0.06	0.04	0.04	0.06	0.08	0.06	0.05	0.06	0.05	0.05
	(0.16)	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)	(0.18)	(0.18)	(0.18)	(0.18)
esg_report_analyzedYes	−0.54*	−0.53*	−0.35	−0.35	−0.35	−0.32	−0.35	−0.37	−0.40	−0.23	−0.20
	(0.25)	(0.25)	(0.27)	(0.27)	(0.27)	(0.27)	(0.28)	(0.28)	(0.28)	(0.30)	(0.29)
treatedYes:esg_report_analyzedYes	0.24	0.24	0.18	0.18	0.19	0.17	0.19	0.20	0.19	0.21	0.16
	(0.34)	(0.34)	(0.34)	(0.34)	(0.34)	(0.34)	(0.34)	(0.34)	(0.35)	(0.35)	(0.34)
esg_driven_inv_recYes		−0.04	0.03	0.03	0.06	0.04	0.05	0.04	0.05	0.07	0.07
		(0.17)	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)	(0.18)	(0.18)	(0.18)
job_type_recManagement			−0.65*	−0.65*	−0.65*	−0.59*	−0.58*	−0.51	−0.53	−0.48	−0.56
			(0.27)	(0.27)	(0.27)	(0.27)	(0.27)	(0.34)	(0.36)	(0.36)	(0.36)
job_type_recSpecialist			−0.41*	−0.41*	−0.40*	−0.36*	−0.35°	−0.26	−0.28	−0.30	−0.32
			(0.18)	(0.18)	(0.18)	(0.18)	(0.18)	(0.22)	(0.22)	(0.23)	(0.22)
esg_importance_recYes					−0.20	−0.17	−0.18	−0.17	−0.18	−0.19	−0.22
					(0.19)	(0.19)	(0.19)	(0.20)	(0.20)	(0.20)	(0.20)
eng_goodYes						−0.35	−0.30	−0.25	−0.26	−0.25	−0.29
						(0.26)	(0.26)	(0.26)	(0.27)	(0.26)	(0.25)
eu_grewYes							−0.18	−0.22	−0.22	−0.24	−0.17
							(0.20)	(0.20)	(0.20)	(0.20)	(0.20)
age4c25-34								−0.23	−0.23	−0.25	−0.21
								(0.19)	(0.20)	(0.19)	(0.20)
age4c35-44								−0.11	−0.12	−0.11	−0.08
								(0.29)	(0.29)	(0.29)	(0.28)
age4c45 and more								−0.21	−0.22	−0.21	−0.13
								(0.40)	(0.40)	(0.39)	(0.39)
firm_report_famYes									0.09	0.24	0.22
									(0.18)	(0.19)	(0.20)
firm_report_analyzedYes										−0.36	−0.39°
										(0.22)	(0.22)
eu_empYes											−1.02*
											(0.47)
R^2	0.04	0.04	0.08	0.08	0.09	0.10	0.10	0.11	0.11	0.12	0.14
Adj. R^2	0.03	0.02	0.06	0.06	0.06	0.06	0.06	0.06	0.05	0.06	0.08
Num. obs.	221	221	221	221	221	221	221	221	221	221	221

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\circ}p < 0.1$

Table 2: Has analyzed an ESG report

	1	2	3	4	5	6	7	8	9	10	11
(Intercept)	4.91^{***}	4.95^{***}	5.08^{***}	5.07^{***}	5.21^{***}	5.44^{***}	5.54^{***}	5.62^{***}	5.57^{***}	5.61^{***}	6.57^{***}
	(0.13)	(0.20)	(0.19)	(0.19)	(0.23)	(0.29)	(0.34)	(0.34)	(0.35)	(0.35)	(0.57)
treatedYes	-0.07	-0.08	-0.14	-0.14	-0.13	-0.09	-0.09	-0.11	-0.10	-0.11	-0.11
	(0.18)	(0.19)	(0.19)	(0.19)	(0.19)	(0.19)	(0.19)	(0.20)	(0.20)	(0.21)	(0.20)
esg_report_readYes	-0.64^{**}	-0.63^{**}	-0.46°	-0.33	-0.33	-0.31	-0.32	-0.29	-0.33	-0.38	-0.35
	(0.23)	(0.23)	(0.24)	(0.29)	(0.29)	(0.29)	(0.30)	(0.30)	(0.31)	(0.31)	(0.31)
treatedYes:esg_report_readYes	0.51°	0.52°	0.54°	0.51°	0.52°	0.49	0.47	0.48	0.48	0.51	0.47
	(0.30)	(0.30)	(0.30)	(0.30)	(0.31)	(0.31)	(0.31)	(0.31)	(0.31)	(0.31)	(0.31)
esg_driven_inv_recYes		-0.06	0.00	0.01	0.04	0.02	0.04	0.03	0.05	0.06	0.06
		(0.17)	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)	(0.18)	(0.18)	(0.18)	(0.18)
job_type_recManagement			-0.66^{*}	-0.66^{*}	-0.66^{*}	-0.61^{*}	-0.59^{*}	-0.52	-0.53	-0.48	-0.56
			(0.28)	(0.28)	(0.28)	(0.28)	(0.28)	(0.36)	(0.36)	(0.37)	(0.37)
job_type_recSpecialist			-0.43^{*}	-0.41^{*}	-0.41^{*}	-0.37°	-0.36°	-0.26	-0.28	-0.31	-0.33
			(0.18)	(0.19)	(0.19)	(0.19)	(0.19)	(0.22)	(0.23)	(0.23)	(0.22)
esg_report_analyzedYes				-0.18	-0.17	-0.16	-0.16	-0.20	-0.21	-0.01	-0.00
				(0.24)	(0.24)	(0.24)	(0.24)	(0.24)	(0.24)	(0.27)	(0.27)
esg_importance_recYes					-0.20	-0.18	-0.19	-0.18	-0.18	-0.20	-0.22
					(0.19)	(0.19)	(0.19)	(0.20)	(0.20)	(0.20)	(0.20)
eng_goodYes						-0.32	-0.28	-0.22	-0.24	-0.24	-0.27
						(0.25)	(0.26)	(0.26)	(0.26)	(0.25)	(0.25)
eu_grewYes							-0.16	-0.20	-0.19	-0.23	-0.15
							(0.20)	(0.20)	(0.20)	(0.20)	(0.20)
age4c25-34								-0.24	-0.23	-0.25	-0.21
								(0.19)	(0.19)	(0.19)	(0.20)
age4c35-44								-0.13	-0.13	-0.12	-0.10
								(0.29)	(0.29)	(0.29)	(0.29)
age4c45 and more								-0.23	-0.23	-0.22	-0.15
								(0.41)	(0.41)	(0.39)	(0.40)
firm_report_famYes									0.11	0.28	0.26
									(0.19)	(0.21)	(0.21)
firm_report_analyzedYes										-0.39°	-0.41°
										(0.23)	(0.22)
eu_empYes											-0.99^{*}
											(0.46)
R^2	0.05	0.05	0.09	0.10	0.10	0.11	0.11	0.12	0.12	0.13	0.15
Adj. R^2	0.04	0.03	0.07	0.07	0.07	0.07	0.07	0.06	0.06	0.07	0.08
Num. obs.	221	221	221	221	221	221	221	221	221	221	221

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\circ}p < 0.1$

Table 3: Has read an ESG report

	1	2	3	4	5	6	7	8	9	10	11
(Intercept)	4.88^{***}	4.93^{***}	5.04^{***}	5.03^{***}	5.17^{***}	5.42^{***}	5.56^{***}	5.63^{***}	5.56^{***}	5.56^{***}	6.55^{***}
	(0.14)	(0.20)	(0.19)	(0.19)	(0.23)	(0.30)	(0.36)	(0.36)	(0.37)	(0.37)	(0.59)
treatedYes	0.02	0.01	−0.01	−0.01	0.00	0.03	0.01	0.00	0.02	0.02	0.02
	(0.18)	(0.18)	(0.18)	(0.18)	(0.18)	(0.18)	(0.19)	(0.19)	(0.19)	(0.19)	(0.19)
firm_report_analyzedYes	-0.59^{*}	-0.58^{*}	-0.42°	−0.35	−0.36	−0.34	−0.38	−0.38	−0.47	−0.47	-0.49°
	(0.23)	(0.23)	(0.23)	(0.27)	(0.27)	(0.27)	(0.27)	(0.27)	(0.29)	(0.29)	(0.28)
treatedYes:firm_report_analyzedYes	0.30	0.30	0.26	0.25	0.26	0.24	0.27	0.25	0.23	0.23	0.20
	(0.30)	(0.31)	(0.30)	(0.30)	(0.31)	(0.31)	(0.31)	(0.32)	(0.32)	(0.32)	(0.32)
esg_driven_inv_recYes		−0.07	0.01	0.02	0.05	0.03	0.05	0.04	0.07	0.07	0.07
		(0.16)	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)	(0.18)	(0.18)	(0.18)	(0.18)
job_type_recManagement			-0.59^{*}	-0.58^{*}	-0.58^{*}	-0.53°	-0.51°	−0.46	−0.49	−0.49	−0.58
			(0.28)	(0.28)	(0.28)	(0.28)	(0.28)	(0.35)	(0.36)	(0.36)	(0.36)
job_type_recSpecialist			-0.42^{*}	-0.40^{*}	-0.39^{*}	-0.35°	-0.34°	−0.26	−0.31	−0.31	−0.33
			(0.17)	(0.18)	(0.18)	(0.18)	(0.18)	(0.22)	(0.23)	(0.23)	(0.22)
esg_report_analyzedYes				−0.12	−0.11	−0.10	−0.10	−0.11	−0.14	−0.14	−0.12
				(0.24)	(0.24)	(0.24)	(0.23)	(0.24)	(0.24)	(0.24)	(0.23)
esg_importance_recYes					−0.20	−0.18	−0.19	−0.19	−0.19	−0.19	−0.22
					(0.19)	(0.19)	(0.19)	(0.20)	(0.20)	(0.20)	(0.20)
eng_goodYes						−0.33	−0.27	−0.21	−0.25	−0.25	−0.29
						(0.25)	(0.25)	(0.25)	(0.25)	(0.25)	(0.25)
eu_grewYes							−0.22	−0.26	−0.25	−0.25	−0.17
							(0.20)	(0.20)	(0.20)	(0.20)	(0.20)
age4c25-34								−0.23	−0.23	−0.23	−0.20
								(0.20)	(0.20)	(0.20)	(0.20)
age4c35-44								−0.08	−0.08	−0.08	−0.06
								(0.29)	(0.29)	(0.29)	(0.29)
age4c45 and more								−0.18	−0.20	−0.20	−0.12
								(0.39)	(0.39)	(0.39)	(0.39)
firm_report_famYes									0.23	0.23	0.22
									(0.19)	(0.19)	(0.19)
eu_empYes											-1.02^{*}
											(0.49)
R^2	0.05	0.05	0.09	0.09	0.10	0.10	0.11	0.12	0.12	0.12	0.14
Adj. R^2	0.04	0.03	0.06	0.06	0.06	0.06	0.07	0.06	0.06	0.06	0.08
Num. obs.	221	221	221	221	221	221	221	221	221	221	221

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\circ}p < 0.1$

Table 4: Has analyzed a firm report

	1	2	3	4	5	6	7	8	9	10	11
(Intercept)	5.09^{***}	5.19^{***}	5.22^{***}	5.17^{***}	5.31^{***}	5.53^{***}	5.65^{***}	5.73^{***}	5.73^{***}	5.76^{***}	6.70^{***}
	(0.15)	(0.20)	(0.21)	(0.21)	(0.25)	(0.30)	(0.35)	(0.36)	(0.36)	(0.36)	(0.61)
treatedYes	-0.36°	-0.38°	-0.41°	-0.40°	-0.39°	-0.34	-0.36°	-0.38°	-0.38°	-0.38°	-0.37°
	(0.21)	(0.21)	(0.21)	(0.21)	(0.21)	(0.21)	(0.21)	(0.22)	(0.22)	(0.22)	(0.22)
firm_report_famYes	-0.69^{***}	-0.69^{***}	-0.45^{*}	-0.32	-0.31	-0.26	-0.29	-0.29	-0.29	-0.14	-0.14
	(0.21)	(0.20)	(0.22)	(0.24)	(0.23)	(0.24)	(0.24)	(0.24)	(0.24)	(0.25)	(0.25)
treatedYes:firm_report_famYes	0.76^{**}	0.77^{**}	0.75^{**}	0.72^{*}	0.73^{**}	0.69^{*}	0.71^{*}	0.72^{*}	0.72^{*}	0.73^{*}	0.69^{*}
	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)	(0.29)	(0.29)	(0.29)	(0.29)
esg_driven_inv_recYes		-0.14	-0.03	0.03	0.06	0.04	0.05	0.05	0.05	0.06	0.06
		(0.16)	(0.16)	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)
job_type_recManagement			-0.69^{*}	-0.66^{*}	-0.66^{*}	-0.62^{*}	-0.60^{*}	-0.52	-0.52	-0.47	-0.55
			(0.28)	(0.28)	(0.28)	(0.28)	(0.28)	(0.35)	(0.35)	(0.36)	(0.36)
job_type_recSpecialist			-0.47^{**}	-0.44^{*}	-0.43^{*}	-0.40^{*}	-0.39^{*}	-0.29	-0.29	-0.31	-0.33
			(0.18)	(0.18)	(0.19)	(0.19)	(0.19)	(0.22)	(0.22)	(0.22)	(0.22)
esg_report_analyzedYes				-0.27	-0.27	-0.26	-0.27	-0.29	-0.29	-0.12	-0.10
				(0.21)	(0.21)	(0.21)	(0.21)	(0.21)	(0.21)	(0.24)	(0.23)
esg_importance_recYes					-0.21	-0.19	-0.20	-0.19	-0.19	-0.20	-0.23
					(0.19)	(0.19)	(0.19)	(0.20)	(0.20)	(0.20)	(0.20)
eng_goodYes						-0.31	-0.27	-0.21	-0.21	-0.20	-0.23
						(0.26)	(0.26)	(0.26)	(0.26)	(0.26)	(0.25)
eu_grewYes							-0.18	-0.23	-0.23	-0.26	-0.18
							(0.20)	(0.20)	(0.20)	(0.20)	(0.20)
age4c25-34								-0.24	-0.24	-0.26	-0.22
								(0.19)	(0.19)	(0.19)	(0.20)
age4c35-44								-0.13	-0.13	-0.12	-0.09
								(0.29)	(0.29)	(0.29)	(0.29)
age4c45 and more								-0.23	-0.23	-0.22	-0.15
								(0.40)	(0.40)	(0.39)	(0.39)
firm_report_analyzedYes										-0.37°	-0.39°
										(0.22)	(0.22)
eu_empYes											-0.97°
											(0.52)
R^2	0.05	0.05	0.10	0.11	0.11	0.12	0.12	0.13	0.13	0.14	0.16
Adj. R^2	0.03	0.03	0.07	0.08	0.08	0.08	0.08	0.07	0.07	0.08	0.10
Num. obs.	221	221	221	221	221	221	221	221	221	221	221

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\circ}p < 0.1$

Table 5: Familiarity with firm reporting

	1	2	3	4	5	6	7	8	9	10	11
(Intercept)	5.04^{***}	5.12^{***}	5.17^{***}	5.13^{***}	5.26^{***}	5.48^{***}	5.59^{***}	5.68^{***}	5.76^{***}	5.77^{***}	6.71^{***}
	(0.15)	(0.20)	(0.21)	(0.22)	(0.25)	(0.29)	(0.34)	(0.35)	(0.39)	(0.38)	(0.62)
treatedYes	-0.31	-0.33	-0.36°	-0.36	-0.34	-0.29	-0.31	-0.33	-0.37°	-0.38°	-0.36°
	(0.21)	(0.21)	(0.22)	(0.22)	(0.21)	(0.21)	(0.22)	(0.22)	(0.22)	(0.22)	(0.22)
firm_report_readYes	-0.63^{**}	-0.63^{**}	-0.39°	-0.26	-0.25	-0.19	-0.21	-0.22	0.52	0.27	0.25
	(0.21)	(0.21)	(0.22)	(0.24)	(0.24)	(0.24)	(0.24)	(0.24)	(1.59)	(1.60)	(1.62)
treatedYes:firm_report_readYes	0.71^{*}	0.71^{*}	0.70^{*}	0.67^{*}	0.68^{*}	0.63^{*}	0.64^{*}	0.66^{*}	0.70^{*}	0.72^{*}	0.68^{*}
	(0.28)	(0.29)	(0.28)	(0.29)	(0.28)	(0.29)	(0.29)	(0.29)	(0.29)	(0.29)	(0.29)
esg_driven_inv_recYes		-0.12	-0.01	0.04	0.07	0.05	0.06	0.06	0.04	0.05	0.06
		(0.16)	(0.16)	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)	(0.18)	(0.17)	(0.18)
job_type_recManagement			-0.70^{*}	-0.67^{*}	-0.67^{*}	-0.63^{*}	-0.62^{*}	-0.53	-0.52	-0.47	-0.55
			(0.28)	(0.28)	(0.27)	(0.28)	(0.28)	(0.35)	(0.36)	(0.36)	(0.36)
job_type_recSpecialist			-0.48^{**}	-0.45^{*}	-0.44^{*}	-0.41^{*}	-0.40^{*}	-0.30	-0.29	-0.31	-0.33
			(0.18)	(0.18)	(0.18)	(0.18)	(0.18)	(0.22)	(0.23)	(0.23)	(0.23)
esg_report_analyzedYes				-0.29	-0.28	-0.28	-0.29	-0.30	-0.30	-0.13	-0.11
				(0.21)	(0.21)	(0.21)	(0.21)	(0.21)	(0.21)	(0.24)	(0.24)
esg_importance_recYes					-0.21	-0.19	-0.19	-0.18	-0.19	-0.20	-0.23
					(0.19)	(0.19)	(0.19)	(0.20)	(0.20)	(0.20)	(0.20)
eng_goodYes						-0.32	-0.27	-0.21	-0.23	-0.21	-0.25
						(0.26)	(0.26)	(0.27)	(0.28)	(0.27)	(0.27)
eu_grewYes							-0.17	-0.22	-0.23	-0.26	-0.18
							(0.20)	(0.20)	(0.21)	(0.20)	(0.20)
age4c25-34								-0.25	-0.23	-0.25	-0.22
								(0.19)	(0.19)	(0.19)	(0.20)
age4c35-44								-0.13	-0.12	-0.12	-0.09
								(0.29)	(0.30)	(0.29)	(0.29)
age4c45 and more								-0.23	-0.22	-0.22	-0.15
								(0.40)	(0.40)	(0.39)	(0.39)
firm_report_famYes									-0.79	-0.41	-0.39
									(1.59)	(1.61)	(1.63)
firm_report_analyzedYes										-0.36	-0.38°
										(0.22)	(0.22)
eu_empYes											-0.97°
											(0.51)
R^2	0.04	0.04	0.09	0.10	0.11	0.11	0.12	0.13	0.13	0.14	0.16
Adj. R^2	0.03	0.03	0.07	0.07	0.07	0.08	0.08	0.07	0.07	0.08	0.09
Num. obs.	221	221	221	221	221	221	221	221	221	221	221

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\circ}p < 0.1$

Table 6: Has read a firm report

	1	2	3	4	5	6	7	8	9	10	11
(Intercept)	4.75^{***}	4.80^{***}	4.98^{***}	5.02^{***}	5.19^{***}	5.41^{***}	5.53^{***}	5.61^{***}	5.57^{***}	5.61^{***}	6.55^{***}
	(0.13)	(0.18)	(0.19)	(0.18)	(0.22)	(0.29)	(0.34)	(0.35)	(0.35)	(0.35)	(0.57)
treatedYes	−0.02	−0.03	−0.07	−0.09	−0.09	−0.06	−0.07	−0.09	−0.08	−0.10	−0.10
	(0.17)	(0.17)	(0.18)	(0.17)	(0.17)	(0.18)	(0.18)	(0.18)	(0.18)	(0.19)	(0.18)
job_esgYes	-0.47°	-0.44°	−0.26	−0.06	−0.09	−0.08	−0.10	−0.10	−0.10	−0.13	−0.12
	(0.26)	(0.26)	(0.25)	(0.29)	(0.29)	(0.29)	(0.29)	(0.30)	(0.30)	(0.30)	(0.30)
treatedYes:job_esgYes	0.65°	0.64°	0.64°	0.66^{*}	0.72^{*}	0.68^{*}	0.70^{*}	0.71^{*}	0.70^{*}	0.76^{*}	0.73^{*}
	(0.35)	(0.35)	(0.33)	(0.32)	(0.33)	(0.33)	(0.33)	(0.34)	(0.34)	(0.35)	(0.34)
esg_driven_inv_recYes		−0.08	−0.00	0.03	0.06	0.05	0.06	0.06	0.07	0.08	0.08
		(0.17)	(0.16)	(0.16)	(0.16)	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)
job_type_recManagement			-0.73^{**}	-0.68^{*}	-0.68^{*}	-0.63^{*}	-0.62^{*}	−0.54	−0.56	−0.50	−0.58
			(0.27)	(0.27)	(0.27)	(0.27)	(0.27)	(0.34)	(0.35)	(0.36)	(0.36)
job_type_recSpecialist			-0.50^{**}	-0.44^{*}	-0.43^{*}	-0.39^{*}	-0.38^{*}	−0.28	−0.30	−0.33	−0.35
			(0.17)	(0.18)	(0.18)	(0.18)	(0.18)	(0.21)	(0.22)	(0.22)	(0.21)
esg_report_analyzedYes				-0.40°	-0.39°	−0.37	−0.38	-0.40°	-0.43°	−0.24	−0.22
				(0.23)	(0.23)	(0.23)	(0.23)	(0.24)	(0.25)	(0.27)	(0.27)
esg_importance_recYes					−0.25	−0.23	−0.24	−0.23	−0.23	−0.25	−0.28
					(0.19)	(0.19)	(0.19)	(0.20)	(0.20)	(0.20)	(0.20)
eng_goodYes						−0.30	−0.25	−0.19	−0.21	−0.20	−0.23
						(0.26)	(0.26)	(0.27)	(0.27)	(0.26)	(0.26)
eu_grewYes							−0.18	−0.23	−0.22	−0.25	−0.18
							(0.19)	(0.20)	(0.20)	(0.20)	(0.19)
age4c25-34								−0.25	−0.25	−0.26	−0.23
								(0.19)	(0.19)	(0.19)	(0.19)
age4c35-44								−0.12	−0.13	−0.12	−0.09
								(0.28)	(0.28)	(0.28)	(0.28)
age4c45 and more								−0.21	−0.22	−0.22	−0.14
								(0.39)	(0.39)	(0.38)	(0.38)
firm_report_famYes									0.09	0.26	0.24
									(0.18)	(0.20)	(0.20)
firm_report_analyzedYes										-0.41°	-0.43°
										(0.23)	(0.23)
eu_empYes											-0.96^{*}
											(0.46)
R^2	0.03	0.03	0.09	0.11	0.11	0.12	0.12	0.13	0.13	0.15	0.16
Adj. R^2	0.02	0.01	0.06	0.08	0.08	0.08	0.08	0.08	0.07	0.08	0.10
Num. obs.	221	221	221	221	221	221	221	221	221	221	221

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\circ}p < 0.1$

Table 7: ESG-related occupation (Expert)

	1	2	3	4	5	6	7	8	9	10	11
(Intercept)	5.15^{***}	5.24^{***}	5.30^{***}	5.26^{***}	5.38^{***}	5.56^{***}	5.68^{***}	5.77^{***}	5.76^{***}	5.79^{***}	6.73^{***}
	(0.16)	(0.21)	(0.22)	(0.22)	(0.25)	(0.29)	(0.34)	(0.35)	(0.35)	(0.35)	(0.61)
treatedYes	-0.56^{**}	-0.58^{**}	-0.64^{**}	-0.62^{**}	-0.60^{**}	-0.55^{*}	-0.57^{*}	-0.60^{**}	-0.60^{**}	-0.60^{**}	-0.59^{*}
	(0.21)	(0.22)	(0.22)	(0.22)	(0.22)	(0.22)	(0.22)	(0.23)	(0.23)	(0.23)	(0.23)
report_famYes	-0.72^{***}	-0.72^{***}	-0.51^{*}	-0.39°	-0.38°	-0.33	-0.36	-0.36	-0.28	-0.30	-0.26
	(0.21)	(0.21)	(0.22)	(0.23)	(0.22)	(0.23)	(0.23)	(0.23)	(0.33)	(0.34)	(0.34)
treatedYes:report_famYes	0.98^{***}	1.00^{***}	1.01^{***}	0.97^{***}	0.97^{***}	0.92^{**}	0.92^{**}	0.96^{**}	0.96^{**}	0.96^{**}	0.92^{**}
	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)	(0.29)	(0.29)	(0.29)	(0.29)
esg_driven_inv_recYes		-0.14	-0.03	0.02	0.05	0.04	0.05	0.04	0.04	0.05	0.05
		(0.16)	(0.16)	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)	(0.18)	(0.17)	(0.18)
job_type_recManagement			-0.74^{**}	-0.69^{*}	-0.69^{*}	-0.65^{*}	-0.64^{*}	-0.52	-0.51	-0.46	-0.54
			(0.27)	(0.27)	(0.27)	(0.28)	(0.28)	(0.35)	(0.35)	(0.36)	(0.36)
job_type_recSpecialist			-0.50^{**}	-0.45^{*}	-0.44^{*}	-0.41^{*}	-0.40^{*}	-0.28	-0.27	-0.29	-0.31
			(0.17)	(0.18)	(0.18)	(0.18)	(0.18)	(0.21)	(0.22)	(0.22)	(0.22)
esg_report_analyzedYes				-0.28	-0.27	-0.26	-0.27	-0.30	-0.29	-0.12	-0.10
				(0.20)	(0.20)	(0.20)	(0.21)	(0.21)	(0.21)	(0.24)	(0.24)
esg_importance_recYes					-0.20	-0.19	-0.19	-0.18	-0.18	-0.19	-0.22
					(0.19)	(0.19)	(0.19)	(0.20)	(0.20)	(0.19)	(0.19)
eng_goodYes						-0.27	-0.23	-0.16	-0.15	-0.15	-0.18
						(0.25)	(0.25)	(0.25)	(0.25)	(0.24)	(0.24)
eu_grewYes							-0.16	-0.21	-0.21	-0.24	-0.17
							(0.19)	(0.20)	(0.20)	(0.20)	(0.20)
age4c25-34								-0.28	-0.28	-0.29	-0.26
								(0.19)	(0.19)	(0.19)	(0.20)
age4c35-44								-0.17	-0.18	-0.16	-0.14
								(0.29)	(0.29)	(0.29)	(0.29)
age4c45 and more								-0.29	-0.30	-0.29	-0.22
								(0.40)	(0.40)	(0.39)	(0.39)
firm_report_famYes									-0.10	0.07	0.04
									(0.31)	(0.32)	(0.32)
firm_report_analyzedYes										-0.36	-0.39°
										(0.22)	(0.22)
eu_empYes											-0.96°
											(0.52)
R^2	0.05	0.05	0.11	0.12	0.12	0.13	0.13	0.14	0.14	0.15	0.17
Adj. R^2	0.04	0.04	0.09	0.09	0.09	0.09	0.09	0.09	0.08	0.09	0.11
Num. obs.	221	221	221	221	221	221	221	221	221	221	221

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\circ}p < 0.1$

Table 8: Familiarity with reporting

9 References

Adams, C. A. (2004). The ethical, social and environmental reportingperformance portrayal gap. *Accounting, Auditing & Accountability Journal*, *17*(5), 731–757. https://doi.org/10.1108/09513570410567791 (cit. on p. 67).

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 (cit. on pp. 64, 65).

Athanasakou, V., & Hussainey, K. (2014). The perceived credibility of forward-looking performance disclosures [publisher: Taylor & Francis]. *Accounting and Business Research*, *44*(3), 227–259. https://doi.org/10.1080/00014788.2013.867403 (cit. on p. 66).

Austin, P. C. (2009). Using the Standardized Difference to Compare the Prevalence of a Binary Variable Between Two Groups in Observational Research [publisher: Taylor & Francis _eprint: https://doi.org/10.1080/03610910902859574]. *Communications in Statistics - Simulation and Computation*, *38*(6), 1228–1234. https://doi.org/10.1080/03610910902859574 (cit. on pp. 77, 78).

Baier, C., Göttsche, M., Hellmann, A., & Schiemann, F. (2021). Too Good To Be True: Influencing Credibility Perceptions with Signaling Reference Explicitness and Assurance Depth [[Online; accessed 2022-03-08]]. *Journal of Business Ethics*. http://link.springer.com/10.1007/s10551-020-04719-7. https://doi.org/10.1007/s10551-020-04719-7 (cit. on pp. 66, 72).

Balluchi, F., Lazzini, A., & Torelli, R. (2021). Credibility of environmental issues in non-financial mandatory disclosure: Measurement and determinants [publisher: Elsevier Ltd]. *Journal of Cleaner Production*, *288*, 125744. https://doi.org/10.1016/j.jclepro.2020.125744 (cit. on p. 66).

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer [publisher: Frontiers]. *Frontiers in Public Health*, *6*. https://www.frontiersin.org/journals/public-health/articles/10.3389/fpubh.2018.00149/full. https://doi.org/10.3389/fpubh.2018.00149 (cit. on p. 73).

Bonner, S. E., & Lewis, B. L. (1990). Determinants of Auditor Expertise [publisher: [Accounting Research Center, Booth School of Business, University of Chicago, Wiley]]. *Journal of Accounting Research*, *28*, 1–20. https://doi.org/10.2307/2491243 (cit. on p. 69).

Bonner, S. E. (2008). Judgment and decision making in accounting [[Online; accessed 2024-08-19]]. *(No Title)*. https://cir.nii.ac.jp/crid/1130000794658041600. https://cir.nii.ac.jp/crid/1130000794658041600 (cit. on p. 68).

Boukes, M., & LaMarre, H. L. (2021). Narrative persuasion by corporate CSR messages: The impact of narrative richness on attitudes and behavioral intentions via character identification, transportation, and message credibility [publisher: Elsevier Inc.]. *Public Relations Review*, *47*(5), 102107. https://doi.org/10.1016/j.pubrev.2021.102107 (cit. on p. 66).

Bouten, L., Everaert, P., Van Liedekerke, L., De 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 (cit. on p. 67).

Carney, M. (2015). Breaking the Tragedy of the Horizon – climate change and financial stability, 1–16 (cit. on p. 67).

Cerulli Associates. (2022). Press Release: Cerulli Anticipates 84$TrillioninWealthTransfersThrough$2045 [[Online; accessed 2024-08-25]]. https://www.cerulli.com/press-releases/cerulli-anticipates-84-trillion-in-wealth-transfers-through-2045 (cit. on p. 76).

Crilly, D., Hansen, M., & Zollo, M. (2016). The grammar of decoupling: A cognitive-linguistic perspective on firms' sustainability claims and stakeholders' interpretation. *Academy of Management Journal*, *59*(2), 705–729. https://doi.org/10.5465/amj.2015.0171 (cit. on p. 69).

Dando, N., & Swift, T. (2003). Transparency and Assurance: Minding the Credibility Gap. *Journal of Business Ethics*, *44*(2-3), 195–200. https://doi.org/10.1023/A:1023351816790 (cit. on pp. 64, 66).

Frederick, D. M., & Libby, R. (1986). Expertise and Auditors' Judgments of conjunctive Events [publisher: [Accounting Research Center, Booth School of Business, University of Chicago, Wiley]]. *Journal of Accounting Research*, *24*(2), 270–290. https://doi.org/10.2307/2491134 (cit. on p. 68).

Gutsche, G., Wetzel, H., & Ziegler, A. (2023). Determinants of individual sustainable investment behavior - A framed field experiment. *Journal of Economic Behavior and Organization*, *209*, 491–508. https://doi.org/10.1016/j.jebo.2023.03.016 (cit. on p. 76).

Hanlon, M., Yeung, K., & Zuo, L. (2022). Behavioral Economics of Accounting: A Review of Archival Research on Individual Decision Makers*. *Contemporary Accounting Research*, *39*(2), 1150–1214. https://doi.org/10.1111/1911-3846.12739 (cit. on p. 68).

Henry, E., & Peytcheva, M. (2020). Joint Effects of Boilerplate and Text Markup on the Judgments of Novice and Experienced Users of Financial Information [publisher: American Accounting Association]. *Behavioral Research in Accounting*, *32*(1), 1–20. https://doi.org/10.2308/bria-52582 (cit. on p. 69).

Hoozée, S., Maussen, S., & Vangronsveld, P. (2019). The impact of readability of corporate social responsibility information on credibility as perceived by generalist versus specialist readers. *Sustainability Accounting, Management and Policy Journal*, *10*(3), 570–591. https://doi.org/10.1108/SAMPJ-03-2018-0056 (cit. on pp. 66, 69, 72).

Libby, R., & Frederick, D. M. (1990). Experience and the Ability to Explain Audit Findings [publisher: [Accounting Research Center, Booth School of Business, University of Chicago, Wiley]]. *Journal of Accounting Research*, *28*(2), 348–367. https://doi.org/10.2307/2491154 (cit. on p. 68).

Libby, R., & Luft, J. (1993). Determinants of judgment performance in accounting settings: Ability, knowledge, motivation, and environment. *Accounting, Organizations and Society*, *18*(5), 425–450. https://doi.org/10.1016/0361-3682(93)90040-D (cit. on pp. 68, 69).

Lock, I., & Schulz-Knappe, C. (2019). Credible corporate social responsibility (CSR) communication predicts legitimacy: Evidence from an experimental study. *Corporate Communications*, *24*(1), 2–20. https://doi.org/10.1108/CCIJ-07-2018-0071 (cit. on p. 66).

Lock, I., & Seele, P. (2016). The credibility of CSR (corporate social responsibility) reports in Europe. Evidence from a quantitative content analysis in 11 countries [publisher: Elsevier Ltd]. *Journal of Cleaner Production*, *122*, 186–200. https://doi.org/10.1016/j.jclepro.2016.02.060 (cit. on p. 66).

Lock, I., & Seele, P. (2017). Measuring Credibility Perceptions in CSR Communication: A Scale Development to Test Readers' Perceived Credibility of CSR Reports. *Management Communication Quarterly*, *31*(4), 584–613. https://doi.org/10.1177/0893318917707592 (cit. on pp. 64, 66, 72, 73).

Luu, E., & Rubio, S. (2023). Millennial managers. *Corporate Governance: An International Review*, (September 2023), 1–24. https://doi.org/10.1111/corg.12564 (cit. on p. 77).

McLaren-Hankin, Y. (2008). 'We expect to report on significant progress in our product pipeline in the coming year': Hedging forward-looking statements in corporate press releases. *Discourse Studies*, *10*(5), 635–654. https://doi.org/10.1177/1461445608094216 (cit. on p. 67).

Muslu, V., Radhakrishnan, S., Subramanyam, K. R., & Lim, D. (2015). Forward-looking md&a disclosures and the information environment. *Management Science*, *61*(5), 931–948. https://doi.org/10.1287/mnsc.2014.1921 (cit. on p. 67).

Reimsbach, D., Hahn, R., & Gürtürk, A. (2018). Integrated Reporting and Assurance of Sustainability Information: An Experimental Study on Professional Investors'

Information Processing. *European Accounting Review*, *27*(3), 559–581. https://doi.org/10.1080/09638180.2016.1273787 (cit. on p. 66).

Reimsbach, D., Schiemann, F., Hahn, R., & Schmiedchen, E. (2020). In the Eyes of the Beholder: Experimental Evidence on the Contested Nature of Materiality in Sustainability Reporting. *Organization and Environment*, *33*(4), 624–651. https://doi.org/10.1177/1086026619875436 (cit. on p. 66).

Schleicher, T., & Walker, M. (2010). Bias in the tone of forward-looking narratives. *Accounting and Business Research*, *40*(4), 371–390. https://doi.org/10.1080/00014788.2010.9995318 (cit. on p. 67).

Senn, S. (1994). Testing for baseline balance in clinical trials [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/sim.4780131703]. *Statistics in Medicine*, *13*(17), 1715–1726. https://doi.org/10.1002/sim.4780131703 (cit. on p. 76).

Sijtsma, K. (2009). On the Use, the Misuse, and the Very Limited Usefulness of Cronbach's Alpha. *Psychometrika*, *74*(1), 107–120. https://doi.org/10.1007/s11336-008-9101-0 (cit. on p. 73).

Stuart, E. A., Lee, B. K., & Leacy, F. P. (2013). Prognostic score–based balance measures can be a useful diagnostic for propensity score methods in comparative effectiveness research [publisher: Elsevier]. *Journal of clinical epidemiology*, *66*(8), S84–S90. https://doi.org/10.1016/j.jclinepi.2013.01.013 (cit. on p. 77).

Tan, H. T., Wang, E. Y., & Zhou, B. (2015). How does readability influence investors' judgments? Consistency of benchmark performance matters. *Accounting Review*, *90*(1), 371–393. https://doi.org/10.2308/accr-50857 (cit. on p. 68).

Torelli, R., Balluchi, F., & Lazzini, A. (2020). Greenwashing and environmental communication: Effects on stakeholders' perceptions. *Business Strategy and the Environment*, *29*(2), 407–421. https://doi.org/10.1002/bse.2373 (cit. on pp. 66, 69).

Tyson, T., & Adams, C. A. (2020). Increasing the scope of assurance research: New lines of inquiry and novel theoretical perspectives. *Sustainability Accounting, Management* *and Policy Journal*, *11*(2), 291–316. https://doi.org/10.1108/SAMPJ-03-2018-0067 (cit. on p. 66).

van Duuren, E., Plantinga, A., & Scholtens, B. (2016). Esg Integration and the Investment Management Process: Fundamental Investing Reinvented [publisher: Springer Netherlands]. *Journal of Business Ethics*, *138*(3), 525–533. https://doi.org/10.1007/s10551-015-2610-8 (cit. on pp. 64, 65).

Chapter 3

Sustainability-Forward-Looking Disclosure and Stock Liquidity: A European Investigation

Imen Derouiche, Melanie Luxembourger, Anke Müßig

Université du Luxembourg, Luxembourg

Abstract

This study examines the effect of sustainability-forward-looking disclosure (SFLD) on stock liquidity. Anecdotal evidence suggests that investors are asking the wrong questions about sustainability since prospective information on sustainability is less likely to be considered by the market. The accounting literature documents the prominent role of forward-looking statements in mitigating information asymmetry between the firm and the market. However, the specificities of SFLD as prospective information make its effect on information asymmetry and, thus, on stock liquidity debatable. On the one hand, as voluntary information, SFLD contributes to mitigating adverse selection and transaction costs, implying lower information asymmetry and, thus, higher stock liquidity. On the other hand, SFLD encompasses information with relatively extreme uncertainty, long-horizon predictability and low verifiability, which is likely to exacerbate information asymmetry and, hence, lower stock liquidity. On the basis of a content analysis approach, we measure the quality of SFLD with reference to the time-horizon precision: the more precise the time- horizon predictability, the higher the SFLD quality. Studying a large sample of 17,763 observations of firms from Western European countries from the 2003-2022 period, we find that better quality of SFLD is associated with higher stock liquidity, meaning that such information contributes to alleviating information asymmetry among traders. Overall, we provide sound evidence consistent with the contribution of SFLD to the corporate information environment despite all the rhetoric about the market perception and expectation of sustainability.

Keywords— Sustainability Disclosure; Stock liquidity; European Firms; Forward- Looking Disclosure

1 Introduction

A large stream of research highlights the importance of integrating non-financial factors such as those related to sustainability in the corporate strategy (Y. Chen et al., 2018; C. Cho et al., 2015; Michelon et al., 2020). Indeed, stakeholders are shown to be more attracted by firms that engage in more sustainable and impact-driven investments (Gamerschlag et al., 2011; Zhang et al., 2021). However, many sustainability experts allege that investors 'are asking the wrong questions on sustainability'. Anecdotal evidence supports the notion that sustainability reports are dominated by historical information, while investors should ask about the corporate strategy and how it addresses risks and opportunities (Adams, 2020), that is, sustainability-forward-looking disclosure (*hereafter* SFLD). The importance of SFLD is increasingly recognized by regulators through mandating disclosure of significant environmental and social risks and strategy such as in Australia and the U.K. or through releasing recommendations on forward-looking disclosures in corporate integrated reporting such as the case of Task force on Climate-related Financial Disclosures (TCFD) (Engle et al., 2020; O'Dwyer and Unerman, 2020).

This research examines the contribution of SFLD to the corporate information environment by investigating the effect of SFLD on stock liquidity. Stock liquidity plays a crucial role in financial markets, serving as a barometer of the efficiency of trade between buyers and sellers. It indicates that both parties are well- informed and can trade without significant price effects (Amihud, 2002; Easley and O'hara, 2004). Information asymmetry among traders appears to be a key factor in stock liquidity, meaning that corporate disclosure policy can influence stock liquidity, as documented by prior accounting literature (R. M. Bushman and Indjejikian, 1995). Diamond and Verrecchia, 1991; Kim and Verrecchia, 1994 suggest that voluntary disclosure reduces information asymmetry between uninformed and informed investors, and thus increases the liquidity of a firm's stock. Leuz and Verrecchia, 2000 show that better disclosure inherent in the voluntary adoption of IFRS by German firms results in greater stock liquidity. Shroff et al., 2013 find that after the passage of the 2005 Securities Offering Reform, firms increased disclosure before seasoned equity offerings, which led to higher liquidity for these firms. Balakrishnan et al., 2014 find that increasing voluntary information by issuing earnings guidance translates into increased liquidity. Schoenfeld, 2017 reports that increased voluntary disclosure of firms joining the S&P 500 index for the first time is associated with increased stock liquidity. In a qualitative study, Graham et al., 2005find that CFOs are more inclined to report an improvement in their firm's stock liquidity when making voluntary disclosure, in particular in small companies. From a different perspective, Christensen et al., 2016 show that increased information requirements implemented by new EU securities regulations bring about a substantial increase in the stock liquidity of EU financial markets.

We extend this strand of literature by investigating the effect of SFLD on stock liquidity. SFLD represents a specific aspect of forward-looking information that differs from earnings forecasts in many respects. First, sustainable investments chiefly result in immediate cash outflows, while the corresponding outcomes are unlikely to be observable or recognized in the short or medium term. The significant time lag between the initiation of sustainable projects and the manifestation of their impacts, extending well beyond conventional financial market timelines, complicates the accurate incorporation of the outcomes of sustainable projects on stock prices Carney, 2015.[1]. Second, as uncommon corporate projects, sustainable investments usually involve high levels of complexity and uncertainty. Most of the corresponding information is relatively vague and has long-horizon predictability, implying potentially high agency and litigation costs. Third, the extreme uncertainty of forward-looking information, in particular, that related to sustainability, requires the implementation of more expert management and corporate governance structures (Raghunathan and Kumar, 2018), which results in additional costs that firms in which sustainability is not among strategic priorities would be unlikely to bear.[2] In addition, given the high technology of sustainable projects, firms are more likely

[1]This is termed by Carney, 2015 as "the tragedy of the horizon".

[2]Greenwashing can be extremely costly for a firm such as Walmart which paid, in 2017, USD 1 million to settle greenwashing claims that misleadingly claimed to sell "biodegradable" or "compostable" plastics or Rusal (the Russian aluminium company) that failed in issuing a green bond raising USD 500 million, despite having launched a green aluminium brand, because of the high risk of green bond coming from an industry known for its energy-intensive and polluting industrial processes (Deschryver and De Mariz, 2020).

to rely on external experts to analyze sustainability aspects and process sustainability-related information, such as external evaluation, second-party opinion, and certification. Fourth, SFLD is chiefly presented in terms of textual narratives rather than being purely quantitative such as earnings forecasts. Thus, managerial earnings forecasts are readily interpreted within a signalling or screening framework, while SFLD would be vulnerable to investors' perceptions and interpretations. Moreover, earnings forecasts are mostly mandated whereas SFLD is voluntary insofar as no precise data are required by the different environmental-related regulations. Fifth, in ambiguous environments, such as firms that are highly committed to sustainability, individuals are more likely to be overoptimistic (Camerer and Lovallo, 1999), which may undermine the quality of SFLD. This is particularly salient for SFLD because it usually provides information on the sustainable perspectives of the company, thus mostly positive information. Indeed, in the absence of a unique or common benchmarking, individuals (i.e., managers) are unlikely to be aware of the gaps in their available information, leading them to be overconfident in their estimates (Malmendier and Tate, 2005).

Given this, the role of SFLD in shaping information asymmetry and, thus, stock liquidity is highly questionable.[3] Two opposite views can be suggested. On the one hand, as voluntary information, SFLD is typically associated with lower adverse selection and decreased transaction costs (Diamond and Verrecchia, 1991; Leuz and Wysocki, 2008), leading to lower information asymmetry and, thus, higher stock liquidity. On the other hand, SFLD typically encompasses information with extreme uncertainty and long-horizon predictability, which could significantly hinder their quality in the eyes of investors, thus potentially leading to lower stock liquidity.

The current research contributes to existing literature in various ways. First, it contributes to the voluntary disclosure literature with a focus on the content analysis framework by studying, for the first time, the information content of forward-looking information related to sustainability. Other studies examining forward-looking information such

[3] generally speaking, the recent debate about excessive disclosures and disclosure overload puts emphasis on the need to distinguish disclosure quality and disclosure quantity (Saha et al., 2019; WBCSD, 2019).

as Bozanic et al., 2018; J. Chen et al., 2019; F. Li, 2010; Muslu et al., 2015 investigate only forward-looking statements in the MD&A section. Second, we add to the content analysis literature by identifying, for the first time, forward- looking statements related to sustainability using a textual narrative approach that combines two keyword lists, i.e., Expectation keyword list and Sustainability keyword list. Both lists are based on existing keyword lists in the literature but are considerably extended and adapted to our context of analysis of annual reports. The quality of SFLD is also, for the first time, gauged with reference to the time-horizon precision of such disclosure. Other studies such as Hope et al., 2016 investigate the quality of textual narratives by focusing on their specificity, i.e., the number of specific entity names, persons, locations, organizations, numeric values, and times and dates. Our quality measure based on time-horizon precision is more appropriate to the nature of sustainability information that would be more credible, that is of higher quality when the time-horizon predictability is precise than when the time-horizon is vague, and even less credible in the absence of any time reference. Besides, existing textual analysis research, as documented by F. Li, 2010, highlights the additional information content present in textual narratives beyond quantitative data. This suggests that information quality can transcend mere numbers. Third, we add to stock liquidity literature by linking it, for the first time, to forward-looking information, more specifically, to the quality of SFLD.

The rest of the paper is organized as follows. In Section 2, we present the theoretical framework and research hypothesis. In Section 3, we describe data and develop the research hypothesis. In Section 4, we present the results of empirical analysis. In Section 5, we conclude the paper.

2 Theoretical Framework and hypothesis

2.1 Theoretical framework

Many studies on corporate disclosure document a sizable proportion of annual report filings containing forward-looking statements on a wide range of topics (Bozanic et al., 2018; F. Li, 2010). Essentially, forward-looking information constitutes a vital component of voluntary disclosure, typically indicating managers' commitment to reducing information asymmetry with market participants and enhancing the transparency of managerial actions for stakeholders (R. Bushman and Smith, 2001; Haggard et al., 2008). Forward-looking information, including SFLD, is by nature particularly useful to investors in their assessment of a firm's prospects because it reflects forecasts through the eyes of management (Krause et al., 2017).

Building on the significance of forward-looking statements, research further reveals that investors value management forecasts, reacting negatively when disclosures do not meet expectations (Baginski et al., 2004; Bozanic et al., 2018), which underscores the credibility and impact of these forecasts on market perceptions. Additionally, the strategic timing of forecasts, particularly by CEOs nearing the end of their tenure, highlights the tactical use of forward-looking information to influence perceptions with minimal personal repercussions, as outcomes are not immediately evident to the market (Cassell et al., 2013).

Moreover, individuals are more inclined to develop thoughts and expectations about the outcomes of their decisions through repeated learning experiences (Morrow and Richards, 1996). Relatedly, managers are likely to develop forecasts, mainly financial forecasts such as earnings, expenditures, and sales, on the basis of historical data. Indeed, if accounting numbers follow a random walk model, the previous year's values are the best predictor of the current values (Board and Walker, 1990). Investors would be aware of such a management prediction process and would be already apt to incorporate such types of managerial forecasts in their assessment of the firm's prospects.

However, investors cannot predict nonfinancial forecasts, particularly those related to sustainability matters, which renders this disclosure particularly informative to them. Compared to other disclosures, the impact of this disclosure on reducing information asymmetry can be attributable to the extreme uncertainty and the long-horizon predictability of such sustainability-related forecasts, making them difficult to verify by stakeholders.

While investor responsiveness underscores the value of managerial forecasts, it also foregrounds the broader debate about excessive disclosures. The disclosure overload puts emphasis on the need to distinguish disclosure quality and disclosure quantity (Saha et al., 2019). The extreme uncertainty and long-horizon predictability of SFLD may substantially hinder their quality in the eyes of users. Indeed, the significant time lag between the initiation of the sustainable projects and the manifestation of their impacts, extending well beyond conventional financial market timelines, complicates the accurate incorporation of the outcomes of sustainable projects on stock prices Carney, 2015.[4]

Moreover, SFLD is left to the discretion of managers who are unlikely to face significant litigation risk for misleading forward-looking statements or omissions in the sustainability report given that: "*courts treat forward-looking ESG statements, commitments, aspirations and intentions as incompatible with the needs and expectations of today's reasonable investor*" (Saad and Strauss, 2020).[5]

The uncertain nature of forward-looking sustainability information demands sophisticated management and governance frameworks, potentially imposing additional costs on firms. Such structures are particularly crucial for companies where sustainability is not a strategic focus (Raghunathan and Kumar, 2018). SFLD benefits from the scrutiny of specialized rating agencies, which is instrumental in elevating their perceived quality among stakeholders (Drempetic et al., 2020; Wong et al., 2021).

However, the proliferation of standard-setting bodies and the diverse methodologies em-

[4]This is termed by Carney, 2015 as "the tragedy of the horizon".

[5]For example, Cazier et al., 2020 report that U.S. federal securities laws provide a safe harbour that effectively shield firms' qualitative forward-looking statements from legal liability.

ployed by sustainability-specialized rating agencies introduce another layer of complexity. The lack of harmonization in evaluating corporate sustainability risks and opportunities fosters confusion and controversy among investors. This confusion is particularly pronounced in environments characterized by high levels of sustainability commitment, where optimism can cloud judgment and lead to overly positive assessments of long-term projects (Camerer and Lovallo, 1999; Choi et al., 2011).

The absence of standardized benchmarking mechanisms exacerbates these challenges, leaving managers potentially unaware of informational gaps. This unawareness can foster overconfidence in their sustainability estimates, complicating investor assessments and further highlighting the intricate dynamics between disclosure practices, investor expectations, and the overarching need for clarity and consistency in sustainability reporting (Malmendier and Tate, 2005).

2.2 Research hypothesis

Investors use management forecasts to gauge a firm's future prospects. This is particularly true for non-financial forecasts related to sustainability, which are less predictable and, therefore, more informative due to their inherent uncertainty and long-horizon implications (De Villiers and Van Staden, 2011).

The credibility of disclosures plays a crucial role in their impact on stock liquidity. Prior research documents that if investors find that disclosures are informative (i.e., credible), this would translate into a positive relation between the incremental information conveyed by such disclosures and stock liquidity. Indeed, incremental disclosures are more likely to provide information that reduces information asymmetry about future cash flows and other economic and non-economic future outcomes, which results in improved decision-making (Biddle et al., 1995; Bowen et al., 1987). However, if disclosures are not credible and window-dressing (i.e., opportunistic), they would be uncorrelated or negatively correlated with stock liquidity.

Regarding CSR disclosures, the forward-looking component could be considered to provide most of the incremental information, which is particularly useful to stakeholders in their decision-making. In this research, we argue that SFLD represents the informative portion of CSR disclosure for at least three reasons. The first reason is the ability to reduce information asymmetry and thus enhance stock liquidity. Historical evidence suggests that forward-looking disclosures, including forecasts and managerial projections, generally offer more significant information content than contemporaneous or historical data. This assertion is supported by findings from Hoskin et al., 1986 and Keung, 2010, who demonstrate that such prospective insights, especially when supplemented with additional forecasts (e.g., sales forecasts), elicit stronger market reactions. This is attributed to the perception that forward-looking information is deemed more informative and reliable by providing a glimpse into the company's future expectations.

The second reason is the textual analysis insights. Research on textual analysis, including studies by Bozanic and Thevenot, 2015; Henry and Leone, 2016; F. Li, 2010; F. Li, 2008, further reinforces the role of forward-looking statements in reducing information asymmetry. These studies highlight the incremental information content of textual narratives within disclosures, particularly those of a forward-looking nature, over quantitative financial data alone. Such narratives are shown to convey new, market-moving information, thereby enhancing market efficiency and improving stock liquidity.

The third reason is that SFLD chiefly represents the unpredicted part of CSR disclosure because historical and contemporaneous aspects of the corporate sustainability strategy are virtually recognized ex-ante by investors through their learning experience of the firm and through timely news releases, but sustainability forecasts are mostly private information, that is more useful to investors in their assessment of the future firm performance. Since the unpredicted component of disclosure is likely to make investors change their expectations regarding information asymmetry in light of new information in the market, that is, have incremental information content relative to the predicted component, we argue that SFLD reduces information asymmetry and enhances stock liquidity.

To further explain the incremental information of SFLD, we provide the example of firms in environmentally sensitive industries, such as oil and gas. These firms are expected to regularly and timely report on the environmental impacts of their activities to highlight their efforts towards 2030 Agenda sustainable development goals (SDG) practices. These firms are also largely covered by specialized environmentalists and medias that continually analyze and screen the damages caused by oil-fossil industry. Since such information channels could already help investors form their opinion about the environmental impacts of the firm' activity, the value of historical environmental disclosure in the annual report would be reduced. Alternatively, investors do not know and cannot predict how these firms

will deal with sustainability externalities, but they expect managers to issue relevant information about their sustainability strategy. These investors could hence discriminate among firms based on sustainability practices if these firms thoroughly explain how they plan to deal with environmental damages by minimizing the possibility of oil spills. Such forward-looking disclosure would thus have an incremental information content compared to the historic information, since it could be more useful to investors in their assessment of the firms' prospects.

Given the potentially important market impact of high-quality forward-looking disclosures, particularly SFLD, we form the following research hypothesis:

Hypothesis. There is a postivie association between SFLD and stock liquidity.

3 Data and research design

3.1 Sample description

The research focuses on firms from seven European countries: Austria, Belgium, France, Germany, Luxembourg, Netherlands, and Switzerland over the 2002- 2022 period. Our period starts in 2002 because Thomson Reuters Asset 4 from Refinitiv Eikon has initiated corporate coverage in 2002. Financial data, corporate governance data, and ESG data were obtained from Refinitiv Eikon. The initial sample encompasses 22,559 observations.

After discarding observations with missing values, the final sample consists of 17,763 firm-year observations of 1,992 unique firms.

Information related to SFLD was obtained from annual reports and separate sustainability reporting if not integrated into the annual report. All these documents were retrieved from Refinitiv Eikon.

3.2 Dependent variable: Stock liquidity

High-quality financial reports are expected to reduce information asymmetry between informed and uninformed market participants (Brown and Hillegeist, 2007). This study employs the bid-ask spread as the primary measure of information asymmetry in financial markets, following the methodology established by Copeland and Galai, 1983 and further supported by subsequent research (Brown and Hillegeist, 2007; Copeland and Galai, 1983; Gajewski and Li, 2015; Glosten and Milgrom, 1985; Khlifi, 2022).

The bid-ask spread is defined as the difference between the highest price a buyer is willing to pay (bid) and the lowest price a seller is willing to accept (ask). This spread is a critical indicator of stock liquidity and serves as a proxy for the level of information asymmetry among market participants. Market makers strategically set bid-ask spreads to balance the revenue from liquidity traders against potential losses from informed traders (Copeland and Galai, 1983).

The rationale for selecting the bid-ask spread as a measure of information asymmetry is grounded in its ability to reflect the cost of immediate execution faced by uninformed traders. This cost arises due to the risk of trading with more informed parties, compelling liquidity providers to demand compensation for this risk (Copeland and Galai, 1983). Thus, a wider bid-ask spread is indicative of higher information asymmetry, as it signifies a greater compensation required by liquidity providers.

The selection of the bid-ask spread as the dependent variable for studying information asymmetry is further justified by its sensitivity to changes in information quality. Literature has demonstrated that the spread adjusts in response to new information entering the market, making it an effective tool for assessing the impact of financial reporting quality on information asymmetry (Dumitrescu and Zakriya, 2022; Siew et al., 2016).

Empirical evidence suggests that enhanced disclosure by firms can significantly reduce information asymmetry between corporate managers and investors, leading to a narrower bid-ask spread (Botosan, 1997; Brown and Hillegeist, 2007). This reduction in the spread is attributable to the decrease in the informational advantage held by informed traders, making the market more efficient and reducing the cost of equity capital.

The integration of SFLD in the reports can play a pivotal role in mitigating information asymmetry. By revealing new sustainability-related market opportunities or concealing them, firms can significantly influence market perceptions and, consequently, the level of information asymmetry reflected in the bid-ask spread.

For instance, in sensitive industries like mining where environmental concerns are crucial, a lack of transparency regarding environmental management or remediation can lead to negative assumptions about a firm's sustainability efforts. This lack of information can widen the bid-ask spread, as the market demands higher compensation for the increased risk associated with greater information asymmetry.

Conversely, proactive disclosures about a firm's environmental strategies reduce market uncertainty by informing investors about the firm's sustainability commitments and environmental risk management. As information asymmetry decreases, the bid-ask spread contracts, signalling improved market efficiency and reduced trading costs associated with asymmetric information.

This dynamic underscores the significance of SFLD in influencing market perceptions and the level of information asymmetry, as reflected in the bid-ask spread. By providing transparency about sustainability practices, firms can effectively reduce the bid-ask spread, indicating a decrease in information asymmetry and an improvement in market efficiency. This, in turn, benefits both the firm and its investors by reducing the cost associated with trading in a market characterized by asymmetric information.

Unlike most prior studies on CSR and its impact on firm performance, which focus on contemporaneous firm performance (e.g., De Villiers and Van Staden, 2011; Freedman and Jaggi, 1988), we examine the firm's long-term expected outcomes of SFLD in terms of stock liquidity and information asymmetry since the effect of such forward-looking information is more likely to manifest in the medium or long run.

Being consistent with prior works (Gajewski and Li, 2015; Khlifi, 2022), the relative spread is measured using the following formula:

$$BAS = \frac{(\mathrm{Ask}_t - \mathrm{Bid}_t)}{(\mathrm{Ask}_t + \mathrm{Bid}_t)/2} * 100$$

where Ask_t is the best ask price at time t, and Bid_t is the best bid price at time t. This formula captures the percentage difference between the ask and bid prices, providing a normalized measure of the bid-ask spread that facilitates comparison across different securities and time periods.

3.3 Measurement of SFLD quality

3.3.1 The process of identifying Forward-Looking Sustainability Disclosure (SFLD)

We use textual analysis to identify SFLD in annual reports. Prior studies, including Bozanic et al., 2018; J. Chen et al., 2019; Muslu et al., 2015, use a library-based approach to identify a list of words that characterize forward-looking statements. Both F. Li, 2010 and Muslu et al., 2015 use two different libraries. We follow a similar library-based approach to identify SFLD by merging the two existing libraries and then extending them in a way that takes into consideration the specificities and the nature of sustainability data.

For the purpose of content analysis, we code the sustainability report into sentences and use them as the unit of analysis. A sustainability sentence is identified as forward-looking if the sentence includes at least one verb implying the future. These phrases are developed based on computational linguistics for identifying future-related sentences. They are completed from our reading of 100 randomly selected sustainability reports. This methodology is to be validated by comparing the automatically generated outputs of 50 randomly selected sustainability reports with a direct reading of the same reports to check whether our methodology is well-specified and powerful.

We now describe in detail our methodology for obtaining SFLD [6]:

- *Step 1: Downloading files:* Our analysis covers firms' annual reports in English. If applicable, separate sustainability reporting are downloaded. Depending on the company, it can be called "ESG reporting", "sustainability reporting", sustainable development reporting", "impact reporting" or "health and safety report".
- *Step 2: Transforming PDF format of Sustainability Reporting into txt format*
- *Step 3: Extracting Sustainability-Related Forward-Looking statements:* We identify a sentence in the sustainability reporting/section as forward looking if the sentence includes at least one forward-looking verb AND at least one term related to Social and Environmental matters. (*Example: The company has committed to maintaining its position as one of the leading European companies with the lowest CO2emissions per kWh produced*).[7] Our analysis is also based on a self-development of a bag of words covering sustainability terms and adapted from Baier et al., 2020 and Sautner et al., 2023.[8]

Corpus subsetting

We subset the initial corpus with all the cleaned sentences and retain only those sentences that contain at least of one of the Sustainability keywords and one of the Expectation keywords. The result is a corpus of forward-looking-sustainable-related sentences. We consider annual reports and separate sustainability reporting if not integrated in the annual report. Since the IFRS does not specify a format or template for annual reports, firms can name sections and headings at their discretion. This makes an automated

[6] We describe in the main text the essential steps. Additional precisions are available in Appendix A
[7] Iberdrola, Statement of Non-Financial Information Sustainability Report 2019.
[8] The two wordlists are provided upon request.

extraction of EU annual report sections difficult. Thus, our method approximates the extraction of forward-looking-sustainable-related sentences in annual report. This does not come without limitations. Our method retains only those sentences that talk about forward-looking-sustainable-related factors and excludes supporting sentences following, which might help explain the factors.

Duplicate detection

However, some duplicates remain because companies might upload both a preliminary and a final version of the English annual report. Thus, for each company and year, we retain only those reports with the highest numbers of total forward-looking-sustainable-related sentences.

3.3.2 Measurement of SFLD quality

Given the specificity of information related to sustainability compared to financial forecasts, we suggest gauging the quality of SFLD with reference to its time-horizon predictability. Indeed, since SFLD basically takes the form of textual narratives, users would be more likely to rely on the forecast horizon to assess the reliability of this disclosure. In this respect, (Muslu et al., 2015) report that the horizon content in forward-looking information makes it more informative to investors in their prediction of future performance.

Since forecasts are generally futuristic and associated with a certain level of uncertainty, we categorize SFLD according to their time-horizon precision. We thus construct three categories of SFLD:

We thus construct three categories of SFLD:

- **Vague range SFLD:** The first category of SFLD includes sentences in which the forecast horizon is not indicated, that is, those without any reference to time. Thus, among Sustainability-Related Forward-Looking sentences, the program tags those not containing numbers or one of following time keywords: “year(s)”, “period(s)”, "decade(s)" (Example: *[...] plan to close the last two coal plants in Spain*).[9]

- **Open-time horizon of SFLD:** The second category of SFLD includes sentences providing a reference to time but without any precision of the forecast time horizon. This category is obtained from the intersection of the SFLD as identified in Step 3 AND one or more keywords of the following list: "next period(s)", "next year(s)", "incoming period(s)", "incoming year(s)", "coming period(s)", "coming year(s)", "upcoming period(s)", "upcoming year(s)", "subsequent period(s)", "subsequent year(s)", "following period(s)", "following year(s)" and "forthcoming year(s)", "Year(s) ahead", "next decade(s)", "incoming decade(s)", "coming decade(s)", "upcoming decade(s)", "subsequent decade(s)", "following decade(s)", "decade(s) ahead", and "forthcoming decade(s)".

- **Precise time-horizon predictability of SFLD:** The third category of SFLD includes sentences providing a reference to a precise point of time. This category is obtained from the intersection of the SFLD as identified in Step 3 AND one or more keywords of the following list: "next X years," "incoming X years," "coming X years," "upcoming X years," "subsequent X years," and "following X years," "forthcoming X years" and "after X year", after a maximum of X years, with X a numerical value which could be written in letters or in numbers, OR if the sentence includes a reference to a year that occurs after the year of filing (such as "2019" or "2020" in a company's sustainability report in 2018).[10] (Examples: "Iberdrola commits to maintaining this indicator 50% below the European average for the sector in the coming 5 years" or "Iberdrola has set the goal of reaching global carbon neutrality by 2050 and expects its emissions intensity in Europe to be practically zero by 2030").[11]

[9] Iberdrola, 2019.

[10] We adapt the word list provided by Muslu et al. (2015) by eliminating periods that are usually employed for financial data such as "next fiscal," "next month," "next quarter" and complement the list from our reading of the randomly selected sample of sustainability reports.

[11] To avoid the erroneous tags of phrases including amounts or percentages whose nominal value occasionally falls within the search range of years, we require that the program discard sentences containing characters in between or before or after the digits (e.g., "euros," "%," ",").

Different from other management expectations, the quality of SFLD cannot be assessed with reference to its ex-ante forecast precision, that is in the short run, such as in Botosan, 1997 and Kent and Ung, 2003. Besides, it is extremely difficult to precisely estimate sustainability impacts which makes it irrelevant to assess SFLD quality exclusively on the basis of numbers reported by management. We also do not tackle the point of the direction of expectations (positive, negative, equal) since most of SFLD indicates perspectives /outlook/projections of the company (thus positive information) rather than risks or threats.

Moreover, SFLD quality does not require that forecasts be made in the short term since many actions and projects related to the environment and climate change cannot be planned unless on a medium or long-term horizon. Based on the previous analysis, we construct the quality index by weighting the three categories of SFLD according to their contribution to information quality and allocating the value 1 to the first category, 2 to the second category, and 3 to the third category.[12] The quality index is the sum of these weighted sub-indices divided by the number of SFLD.

$$\text{SFLD QUAL} = \frac{\sum \text{Weighted values of SFLD}}{\text{Number of SFLD}}$$

Where: SFLD is sentences including words from both sustainability and forward-looking list words; $\sum$ Weighted values of SFLD is the sum (first category SFLD $\times$ 1 + second category SFLD $\times$ 2 + third category SFLD $\times$ 3).

3.4 Control variables

SIZE

Following prior literature (Scott, 1994), we use the log of sales to proxy the size. Past literature has shown a negative link between size and the bid-ask spread and a positive relation between size and volume.

[12]For example, Deloitte Consulting (2003) found that Enron's corporate communications became increasingly vague and ambiguous as the firm's financial situation began to deteriorate.

ROA

Jovanovic, 1982 suggests that higher-quality sellers may be more forthcoming with information. We use ROA, the return on assets calculated by dividing earnings before extraordinary items by total assets, as a proxy for profitability. More profitable firms may disclose more information, and higher ROA might signal a strong financial position, enhancing investor confidence and potentially narrowing the spread due to increased demand.

AGE

Younger firms typically face higher information asymmetry because they have shorter histories and potentially less information available to the public. This lack of information can lead to greater uncertainty among investors about the firm's value and prospects, resulting in a wider bid-ask spread. This variable is measured as the natural logarithm of number of years since establishment

VOLATILITY

Return volatility is the annual standard deviation of daily returns. Botosan and Plumlee, 2002 and S. Cho et al., 2013 have demonstrated a significant correlation between volatility and bid-ask spread, indicating that higher volatility often leads to wider spreads.

PRICE

Price level allows for the effect of discreteness to be controlled. Stocks with low price levels tend to be new, smaller in size and above all riskier. This additional risk leads to an enlargement of the bid–ask spread. Stoll, 2000 has proven that the relative spread is negatively related to the price level in logarithm. The stock price is therefore measured as the logarithm of the average closing stock price.

Table 1 summarizes the control variables considered.

[Insert Table 3.1 about here]

4 Data description and empirical analysis

4.1 Descriptive statistics and correlation analysis

Table 2 presents a summary of the dataset categorized by country. It lists the total number of observations and firms surveyed in various European countries. Germany leads with the highest number of observations, 5,704, and the most firms, 654. In contrast, Luxembourg, despite having fewer observations (605), shows a relatively high number of firms (99) compared to its size.

[Insert Table 3.2 about here]

Table 3 presents the pairwise correlations among the key variables: BAS, SFLD_QUAL, SIZE, ROA, AGE, VOLATILITY, and PRICE. The correlation coefficients indicate the linear relationship between each pair of variables, with the level of statistical significance denoted by asterisks.

[Insert Table 3.3 about here]

The strong negative correlations indicate that larger, more profitable, and older firms tend to experience higher liquidity, while volatility is positively correlated with the bid-ask spread, reflecting higher information asymmetry and reduced liquidity for more volatile stocks.

4.2 Quality of SFLD and stock liquidity

Table 4 reports the results of regression estimating the effect of SFLD quality on stock liquidity. The dependent variable in model (1) is the Bid-ask spread, which we estimate using the pooled OLS model. Results show a negative and statistically significant coefficient of SFLD, suggesting that firms having better-quality SFLD have higher stock liquidity (i.e., a narrower spread) and, thus, lower information asymmetry. This demonstrates that SFLD contributes to improving the corporate information environment between buyers and sellers, thus reducing the risk premium demanded by liquidity providers.

Turning to the effect of control variables, we find that firm size correlates positively with the bid-ask spread, indicating that larger firms tend to have larger spreads. Firm age is negatively associated with the bid-ask spread, aligning with the expectation that older, more established firms likely have more stable trading patterns and possibly more transparency, leading to reduced information asymmetry. Regarding firm volatility, the correlation with stock liquidity is found to be positive, indicating that higher volatility, reflecting greater uncertainty about the firm's future performance, is associated with wider spreads. This relationship highlights the market's response to risk, with liquidity providers demanding higher compensation for trading in more volatile stocks.

We include industry-fixed effects because industry is potentially highly relevant to the effect of SFLD quality on stock liquidity.

The inclusion of country-fixed effects in regression models helps control unobserved heterogeneity at the national level that may influence the dependent variable, the bid-ask spread. In the context of report quality, models without country effects may not fully capture the unique attributes and conditions inherent to each country that could impact information asymmetry. By incorporating country-fixed effects, the analysis adjusts for these unobserved national characteristics, leading to a more refined understanding of the actual influence of report quality on the bid-ask spread.

The dependent variable in model (2) is the Bid-Ask spread which we estimate using a firm fixed effect model. The inclusion of both firm fixed effects and year effects in the regression models controls for unobserved heterogeneity both within firms and over years, leading to more reliable and generalizable results. The inclusion of firm fixed effects controls for temporal variations that could influence the market environment and the bid-ask spread, such as macroeconomic changes, market-wide trends, and regulatory shifts that occur over time. Adjusting for these firm-specific factors, the models continue to show a significant negative coefficient for SFLD quality, affirming the robustness of the relationship between report quality and information asymmetry. This persistence highlights the importance of report quality as a determinant of the bid-ask spread, independent of firm-specific and temporal fluctuations. It suggests that the influence of reporting quality on reducing information asymmetry is not merely a reflection of firm-specific events or conditions but is a stable and enduring factor.

[Insert Table 3.4 about here]

4.3 Robustness checks

To check the robustness of our results, we use an alternative measurement of stock liquidity, namely the Volume Coefficient of Variation (VCV) which is used as a measure of the variability of trade volumes over time (Lof and van Bommel, 2023). A higher VCV indicates more variability in trading volumes, which may be associated with higher information asymmetry or uncertainty about the stock's value, and thus lower stock liquidity. The bid-ask spread is a direct measure of information asymmetry and liquidity in the market, whereas VCV is an indirect measure that captures variability and potential uncertainty. While both are related to information asymmetry, they may capture different aspects of market behavior. For instance, the bid-ask spread is a static measure at a point in time, while VCV reflects dynamic behavior over a period. Similar to the approach adopted for the variable bid-ask spread, and because a higher value of this measure corresponds to a lower level of liquidity.

Results of estimating VCV are reported in Table 5. Model (1) uses pooled OLS estimation whereas Model (2) uses a firm fixed-effect estimation. Results show that, in both models, the coefficient for the quality of SFLD is negative and statistically significant. This suggests that higher-quality SFLD is associated with less variability in trading volume, that is, lower information asymmetry. This aligns with the idea that better-quality information can lead to more stable investor behavior, as investors have clearer expectations about the firm's sustainability and future performance.

[Insert Table 3.5 about here]

5 Conclusion

This study investigates the quality of sustainability-forward-looking disclosure (SFLD) and its impact on information asymmetry in financial markets. By focusing on the impact of the quality of SFLD on stock liquidity and reduction of information asymmetry, this research has illuminated the importance of nonfinancial factors in corporate reporting and strategy. The study has extended existing methodologies by introducing a novel approach to evaluate the quality of SFLD through time-horizon precision, thereby providing a more nuanced understanding of how such disclosures influence market perceptions and thus stock liquidity.

Studying a large sample of 17,763 observations of firms from Western European countries from the 2003-2022 period, we find that better quality of SFLD is associated with higher stock liquidity, meaning that such information contributes to alleviating information asymmetry among traders. The research demonstrates that even in the face of regulatory discretion and the voluntary nature of sustainability disclosures, there is a discernible and significant impact on market efficiency when firms provide relatively clearer and more credible forward-looking information about their sustainability initiatives.

Moreover, the incorporation of both country and year fixed effects in the regression models has strengthened the robustness of the findings, ensuring that the observed effects are not confounded by unobserved heterogeneity or temporal macroeconomic factors. The persistence of the negative relationship between the quality of SFLD and the bid-ask spread, even after controlling for such effects, is a testament to the enduring value of high-quality sustainability disclosures.

The insights derived from this study have practical implications for managers and policymakers, suggesting that improving the quality of sustainability disclosures can be a strategic tool to enhance market transparency and reduce the cost of equity capital. As firms and regulators strive to meet the growing demands of stakeholders for meaningful and actionable sustainability information, the findings of this research highlight the pathway for achieving more efficient and informed financial markets through high-quality SFLD.

The research presented in this paper has taken significant strides in understanding the value relevance of SFLD and its impact on reducing information asymmetry. However, several research gaps remain, presenting opportunities for future studies to build upon the findings of this work:

Geographic Expansion: While this study focuses on European firms, future research could explore the impact of SFLD in other regions. Comparative studies between developed and emerging markets could yield insights into how different economic and regulatory environments influence the effectiveness of SFLD.

Qualitative Approaches: Qualitative investigations into how managers and investors perceive and use SFLD could complement the quantitative findings of this paper, providing a more in-depth understanding of the mechanisms behind the observed relationships.

6 Figures and Tables

Table 3.1: Description of Variables

Variable	Description
BAS	Relative Bid-ask Spread, measured as: $$\text{BAS} = \left(\frac{\text{Ask}_t - \text{Bid}_t}{(\text{Ask}_t + \text{Bid}_t)/2}\right) \times 100$$ Because a higher value of this measure corresponds to a lower level of liquidity, for ease of interpretation, we multiply it by -1 in the empirical tests.
SFLD QUAL	SFLD quality is measured as follows: $$\text{SFLD QUAL} = \frac{\sum \text{Weighted values of SFLD}}{\text{Number of SFLD}}$$ Where: SFLD is sentences including words from both sustainability and forward-looking list words; and $\sum$ Weighted values of SFLD is the sum (first category SFLD×1 + second category SFLD×2 + third category SFLD×3).
SIZE	Firm size, measured as the natural logarithm of total sales
ROA	Firm profitability, measured as the ratio Return on Assets
AGE	Firm age, measured as the log of number of years since establishment
VOLATILITY	Return volatility, measured as the annual standard deviation of daily returns
PRICE	Average daily closing stock price, measured as the log of the average closing price

Table 3.2: Descriptive statistics

	Panel A. Distribution by country			
Country	Nb.obs	Freq	Nb.firms	Freq
Austria	976	5%	97	5%
Belgium	1,641	9%	161	8%
Switzerland	3,374	19%	313	16%
Germany	5,704	32%	654	33%
France	3,525	20%	423	21%
Luxembourg	605	3%	99	5%
Netherlands	1,938	11%	245	12%
Total sample	**17,763**	**100%**	**1,992**	**100%**

	Panel B. Summary statistics						
	mean	sd	p10	p25	p50	p75	p90
BAS	2.6167	1.4605	1.8876	1.7248	2.3994	3.2872	4.3968
SFLD_QUAL	1.3643	0.2289	1.0741	1.2000	1.3556	1.5000	1.6429
SIZE	20.2223	2.1461	17.1936	18.7203	20.1749	21.8249	23.2803
ROA	2.2216	8.3275	-7.4287	0.3801	3.3376	6.6109	10.9534
AGE	35.6163	33.8212	7	13	23	45	90
VOLATILITY	2.5704	1.9937	1.1958	1.5480	2.1083	2.9403	4.2807
PRICE	3.1535	1.5742	1.2602	2.2429	3.1605	4.4046	4.8587

Table 3.3: Correlations

	BAS	SFLD_QUAL	SIZE	ROA	AGE	VOLATILITY	PRICE
BAS	1.0000***						
SFLD_QUAL	-0.0772***	1.0000***					
SIZE	-0.0845***	0.1082***	1.0000***				
ROA	-0.3320***	0.0187*	0.2877***	1.0000***			
AGE	-0.1252***	0.0781***	0.3178***	0.0822***	1.0000***		
VOLATILITY	0.4775***	-0.0887***	-0.2793***	-0.3439***	-0.1258***	1.0000***	
PRICE	-0.2412***	0.0674***	0.3280***	0.2853***	0.2003***	-0.2952***	1.0000***

This table reports the correlation analysis of the variables used. All variables are described in Appendix A. Asterisks indicate significance levels: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table 3.4: SFLD quality and stock liquidity: Main analysis

VARIABLES	**(1) Pooled OLS**	**(2) Fixed Effects**
SFLD_QUAL	-0.110***	-0.095***
	(0.042)	(0.031)
SIZE	0.086***	0.132***
	(0.005)	(0.015)
ROA	-0.033***	-0.012***
	(0.001)	(0.001)
AGE	-0.003***	-0.037***
	(0.000)	(0.002)
VOLATILITY	0.250***	0.305***
	(0.005)	(0.004)
PRICE	-0.096***	-0.083***
	(0.007)	(0.010)
Year dummies	Yes	Yes
Industry dummies	Yes	No
Country dummies	Yes	No
Firm-fixed effects	No	Yes
Constant	1.302***	0.974***
	(0.145)	(0.289)
Observations	17,763	17,763
R-squared	0.374	0.462
Number of firms	1,992	1,992

This table reports the results of estimating the effect of SFLD quality on stock liquidity. Dependent variable is Bid-ask spread. Pooled OLS estimation is used in Model (1). Firm fixed-effects estimation is used in Model (2). All variables are described in Appendix A. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 3.5: SFLD quality and stock liquidity: Robustness checks

VARIABLES	**(1) Pooled OLS**	**(2) Fixed effects**
SFLD_QUAL	-0.184***	-0.076*
	(0.052)	(0.042)
SIZE	-0.233***	-0.129***
	(0.007)	(0.021)
ROA	-0.003**	-0.003**
	(0.001)	(0.001)
AGE	0.001*	-0.019***
	(0.000)	(0.003)
VOLATILITY	0.087***	0.025***
	(0.006)	(0.006)
PRICE	-0.001	-0.107***
	(0.008)	(0.014)
Constant	6.074***	5.282***
	(0.179)	(0.397)
Year dummies	Yes	Yes
Industry dummies	Yes	No
Country dummies	Yes	No
Firm-fixed effects	No	Yes
Observations	17,714	17,714
R-squared	0.166	0.033

This table reports the results of robustness checks for estimating the effect of SFLD quality on stock liquidity. Dependent variable is the Volume Coefficient of Variation (VCV) as a measure of the variability of trade volumes over time. Pooled OLS estimation is used in Model (1). Firm fixed-effects estimation is used in Model (2). All variables are described in Appendix A. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

7 Appendices

Appendix A - Collection of annual reports of EU companies Fetching annual reports and supporting information

The annual reports are batch downloaded from the Thomson Reuters Eikon desktop app. In addition to the annual reports, we download for each company an accompanying Excel file from the Thomson Reuters Eikon Desktop app that contains the company name, the filing date, the receipt date, and the document date of the downloaded annual reports. These Excel files enable us to allocate the PDF/.txt files to the right fiscal year. When we batch download the annual reports, the name of the files contains the company name and the filing date of the annual report. Thus, the accompanying Excel files allow us to assign a document date, which is the fiscal year end date of the annual report, to the file name of the document. Subsequently, we match the textual measures calculated to the financial data according to the company identifier and the fiscal year end.

Sentence cleaning

The PDF documents are converted into plain text .txt format. Some PDF files are not converted properly (i.e., if the text consists partly or entirely of symbols). Following Lang and Stice-Lawrence (2015), we clean sentences to deal with the improperly converted parts of the text and to ensure that tables are not taken into consideration. First, we tokenize the text into sentences. Next, we keep only those sentences that contain at least 50% letters. We measure the percentage of letters as the ratio of letters in a sentences to the sum of the number of letters and digits. We do this to ensure that the sentences we retain are actual sentences and not tables wrongfully tokenized as a sentence.

Then, following Lang and Stice-Lawrence (2015), we retain only those sentences where less than 20% of the characters are non-alphanumeric. We calculate the percentage of non-alphanumeric characters as one minus the percentage of alphanumeric characters. The percentage of alphanumeric characters is the sum of the number of letters and digits per sentence divided by the total number of characters per sentence. This is done to remove all sentences that consist mostly of symbols, which suggests that the text was not converted properly.

With the aim of excluding page headings or page numberings recorded as sentences, we add the condition that the sentences have to contain at least 30 letters.

Finally, to ensure that no tables were counted as sentences, we add the condition that the total count of digits per sentence must not exceed 30.

Only those sentences that fulfill all four conditions are retained in the corpus; all other sentences are removed.

8 References

Adams, C. (2020). Sustainability reporting and value creation. *Social and Environmental Accountability Journal*, *40*(3), 191–197. https://doi.org/10.1080/0969160X.2020.1837643 (cit. on p. 132).

Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets*, *5*(1), 31–56 (cit. on p. 132).

Baginski, S., Hassell, J., & Kimbrough, M. (2004). Why do managers explain their earnings forecasts? *Journal of Accounting Research*, *42*(1), 1–29. https://doi.org/10.1111/j.1475-679X.2004.00127.x (cit. on p. 136).

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 (cit. on p. 144).

Balakrishnan, K., Billings, M. B., Kelly, B., & Ljungqvist, A. (2014). Shaping liquidity: On the causal effects of voluntary disclosure. *the Journal of Finance*, *69*(5), 2237–2278 (cit. on p. 133).

Biddle, G. C., Seow, G. S., & Siegel, A. F. (1995). Relative versus incremental information content. *Contemporary accounting research*, *12*(1), 1–23 (cit. on p. 138).

Board, J., & Walker, M. (1990). Intertemporal and cross-sectional variation in the association between unexpected accounting rates of return and abnormal returns. *Journal of Accounting Research*, 182–192 (cit. on p. 136).

Botosan, C. (1997). Disclosure level and the cost of equity capital. *The Accounting Review*, *72*(3), 323–349 (cit. on pp. 142, 147).

Botosan, C., & Plumlee, M. (2002). A re-examination of disclosure level and the expected cost of equity capital. *Journal of Accounting Research*, *40*(1), 21–40. https://doi.org/10.1111/1475-679X.00037 (cit. on p. 148).

Bowen, R. M., Burgstahler, D., & Daley, L. A. (1987). The incremental information content of accrual versus cash flows. *Accounting Review*, 723–747 (cit. on p. 138).

Bozanic, Z., Roulstone, D., & Van Buskirk, A. (2018). Management earnings forecasts and other forward-looking statements. *Journal of Accounting and Economics*, *65*(1), 1–20. https://doi.org/10.1016/j.jacceco.2017.11.008 (cit. on pp. 135, 136, 143).

Bozanic, Z., & Thevenot, M. (2015). Qualitative disclosure and changes in sell-side financial analysts' information environment. *Contemporary Accounting Research*, *32*(4), 1595–1616 (cit. on p. 139).

Brown, S., & Hillegeist, S. (2007). How disclosure quality affects the level of information asymmetry. *Review of Accounting Studies*, *12*(2–3), 443–477. https://doi.org/10.1007/s11142-007-9032-5 (cit. on pp. 141, 142).

Bushman, R., & Smith, A. (2001). Financial accounting information and corporate governance. *Journal of Accounting and Economics*, *32*(1–3), 237–333 (cit. on p. 136).

Bushman, R. M., & Indjejikian, R. J. (1995). Voluntary disclosures and the trading behavior of corporate insiders. *Journal of Accounting Research*, *33*(2), 293–316 (cit. on p. 132).

Camerer, C., & Lovallo, D. (1999). Overconfidence and excess entry: An experimental approach. *American economic review*, *89*(1), 306–318 (cit. on pp. 134, 138).

Carney, M. (2015). Breaking the tragedy of the horizon – climate change and financial stability [No. September, pp. 1–16]. (Cit. on pp. 133, 137).

Cassell, C., Huang, S., & Sanchez, J. (2013). Forecasting without consequence? evidence on the properties of retiring ceos' forecasts of future earnings. *The Accounting Review*, *88*(6), 1909–1937 (cit. on p. 136).

Cazier, R. A., Merkley, K. J., & Treu, J. S. (2020). When are firms sued for qualitative disclosures? implications of the safe harbor for forward-looking statements. *The Accounting Review*, *95*(1), 31–55 (cit. on p. 137).

Chen, J., Kama, I., & Lehavy, R. (2019). A contextual analysis of the impact of managerial expectations on asymmetric cost behavior. *Review of Accounting Studies*, *24*, 665–693 (cit. on pp. 135, 143).

Chen, Y., Hung, M., & Wang, Y. (2018). The effect of mandatory csr disclosure on firm profitability and social externalities: Evidence from china. *Journal of Accounting and Economics*, *65*(1), 169–190. https://doi.org/10.1016/j.jacceco.2017.11.009 (cit. on p. 132).

Cho, C., Michelon, G., Patten, D., & Roberts, R. (2015). Csr disclosure: The more things change...? *Accounting, Auditing & Accountability Journal*, *28*(1), 14–35 (cit. on p. 132).

Cho, S., Lee, C., & Pfeiffer, R. (2013). Corporate social responsibility performance and information asymmetry. *Journal of Accounting and Public Policy*, *32*(1), 71–83. https://doi.org/10.1016/j.jaccpubpol.2012.10.005 (cit. on p. 148).

Choi, J.-H., Myers, L. A., Zang, Y., & Ziebart, D. A. (2011). Do management eps forecasts allow returns to reflect future earnings? implications for the continuation of management's quarterly earnings guidance. *Review of Accounting Studies*, *16*, 143–182 (cit. on p. 138).

Christensen, H. B., Hail, L., & Leuz, C. (2016). Capital-market effects of securities regulation: Prior conditions, implementation, and enforcement. *The Review of Financial Studies*, *29*(11), 2885–2924 (cit. on p. 133).

Copeland, T., & Galai, D. (1983). Information effects on the bid-ask spread. *The Journal of Finance*, *38*(5), 1457–1469. https://doi.org/10.1111/j.1540-6261.1983.tb03834.x (cit. on p. 141).

De Villiers, C., & Van Staden, C. J. (2011). Where firms choose to disclose voluntary environmental information. *Journal of Accounting and Public Policy*, *30*(6), 504–525 (cit. on pp. 138, 143).

Deschryver, P., & De Mariz, F. (2020). What future for the green bond market? how can policymakers, companies, and investors unlock the potential of the green bond market? *Journal of risk and Financial Management*, *13*(3), 61 (cit. on p. 133).

Diamond, D. W., & Verrecchia, R. E. (1991). Disclosure, liquidity, and the cost of capital. *The journal of Finance*, *46*(4), 1325–1359 (cit. on pp. 132, 134).

Drempetic, 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*, 333–360 (cit. on p. 137).

Dumitrescu, A., & Zakriya, M. (2022). Governance, information flow, and stock returns. *Journal of Corporate Finance*, *72*, 102168. https://doi.org/10.1016/j.jcorpfin.2022.102168 (cit. on p. 142).

Easley, D., & O'hara, M. (2004). Information and the cost of capital. *The journal of finance*, *59*(4), 1553–1583 (cit. on p. 132).

Engle, R., Giglio, S., Kelly, B., Lee, H., & Stroebel, J. (2020). Hedging climate change news. *The Review of Financial Studies*, *33*(3), 1184–1216 (cit. on p. 132).

Freedman, M., & Jaggi, B. (1988). An analysis of the association between pollution disclosure and economic performance. *Accounting, Auditing & Accountability Journal*, *1*(2), 43–58 (cit. on p. 143).

Gajewski, J., & Li, L. (2015). Can internet-based disclosure reduce information asymmetry? *Advances in Accounting*, *31*(1), 115–124. https://doi.org/10.1016/j.adiac.2015.03.013 (cit. on pp. 141, 143).

Gamerschlag, R., Möller, K., & Verbeeten, F. (2011). Determinants of voluntary csr disclosure: Empirical evidence from germany. *Review of Managerial Science*, *5*, 233–262 (cit. on p. 132).

Glosten, L., & Milgrom, P. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, *14*(1), 71–100. https://doi.org/10.1016/0304-405X(85)90044-3 (cit. on p. 141).

Graham, J. R., Harvey, C. R., & Rajgopal, S. (2005). The economic implications of corporate financial reporting. *Journal of accounting and economics*, *40*(1-3), 3–73 (cit. on p. 133).

Haggard, K., Martin, X., & Pereira, R. (2008). Does voluntary disclosure improve stock price informativeness? *Financial Management*, *37*(4), 747–768 (cit. on p. 136).

Henry, E., & Leone, A. J. (2016). Measuring qualitative information in capital markets research: Comparison of alternative methodologies to measure disclosure tone. *The Accounting Review*, *91*(1), 153–178 (cit. on p. 139).

Hope, O., Hu, D., & Lu, H. (2016). The benefits of specific risk-factor disclosures. *Review of Accounting Studies*, *21*. https://doi.org/10.1007/s11142-016-9371-1 (cit. on p. 135).

Hoskin, R. E., Hughes, J. S., & Ricks, W. E. (1986). Evidence on the incremental information content of additional firm disclosures made concurrently with earnings. *Journal of accounting research*, 1–32 (cit. on p. 139).

Iberdrola. (2019). Statement of non-financial information sustainability report 2019 [Accessed: 2025-01-15]. *Iberdrola*. https://ceowatermandate.org/wp-content/uploads/2020/06/IB_Sustainability_Report_2019.pdf (cit. on p. 146).

Jovanovic, B. (1982). Truthful disclosure of information. *The Bell Journal of Economics*, *13*(1), 36. https://doi.org/10.2307/3003428 (cit. on p. 148).

Kent, P., & Ung, K. (2003). Voluntary disclosure of forward-looking earnings information in australia. *Australian Journal of Management*, *28*(3), 273–285 (cit. on p. 147).

Keung, E. C. (2010). Do supplementary sales forecasts increase the credibility of financial analysts' earnings forecasts? *The Accounting Review*, *85*(6), 2047–2074 (cit. on p. 139).

Khlifi, F. (2022). Web-based financial reporting, social media and information asymmetry: The case of saudi arabia. *Journal of Financial Reporting and Accounting*, *20*(5), 994–1020. https://doi.org/10.1108/JFRA-01-2021-0008 (cit. on pp. 141, 143).

Kim, O., & Verrecchia, R. E. (1994). Market liquidity and volume around earnings announcements. *Journal of accounting and economics*, *17*(1-2), 41–67 (cit. on p. 132).

Krause, J., Sellhorn, T., & Ahmed, K. (2017). Extreme uncertainty and forward-looking disclosure properties. *Abacus*, *53*(2), 240–272. https://doi.org/10.1111/abac.12100 (cit. on p. 136).

Leuz, C., & Verrecchia, R. E. (2000). The economic consequences of increased disclosure. *Journal of accounting research*, 91–124 (cit. on p. 132).

Leuz, C., & Wysocki, P. D. (2008). Economic consequences of financial reporting and disclosure regulation: A review and suggestions for future research. *Available at SSRN 1105398* (cit. on p. 134).

Li, F. (2010). The information content of forward-looking statements in corporate filings – a naïve bayesian machine learning approach. *Journal of Accounting Research*, *48*(5), 1049 (cit. on pp. 135, 136, 139, 143).

Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and economics*, *45*(2-3), 221–247 (cit. on p. 139).

Lof, M., & van Bommel, J. (2023). Asymmetric information and the distribution of trading volume. *Journal of Corporate Finance*, *82*, 102464 (cit. on p. 151).

Malmendier, U., & Tate, G. (2005). Ceo overconfidence and corporate investment. *The journal of finance*, *60*(6), 2661–2700 (cit. on pp. 134, 138).

Michelon, G., Rodrigue, M., & Trevisan, E. (2020). The marketization of a social movement: Activists, shareholders and csr disclosure. *Accounting, Organizations and Society*, *80*, 101074. https://doi.org/10.1016/j.aos.2019.101074 (cit. on p. 132).

Morrow, V., & Richards, M. (1996). The ethics of social research with children: An overview 1. *Children & society*, *10*(2), 90–105 (cit. on p. 136).

Muslu, V., Radhakrishnan, S., Subramanyam, K., & Lim, D. (2015). Forward-looking md&a disclosures and the information environment. *Management Science*, *61*(5), 931–948 (cit. on pp. 135, 143, 145).

O'Dwyer, B., & Unerman, J. (2020). Shifting the focus of sustainability accounting from impacts to risks and dependencies: Researching the transformative potential of tcfd reporting. *Accounting, Auditing & Accountability Journal*, *33*(5), 1113–1141 (cit. on p. 132).

Raghunathan, M., & Kumar, A. (2018). Csr committee composition and possible impacts. *Journal of Contemporary Research in Management*, *13*(4), 9–27 (cit. on pp. 133, 137).

Saad, A., & Strauss, D. (2020). A new "reasonable investor" and changing frontiers of materiality: Increasing investor reliance on esg disclosures and implications for securities litigation. *Berkeley Bus. LJ*, *17*, 391 (cit. on p. 137).

Saha, A., Morris, R. D., & Kang, H. (2019). Disclosure overload? an empirical analysis of international financial reporting standards disclosure requirements. *Abacus*, *55*(1), 205–236 (cit. on pp. 134, 137).

Sautner, Z., Van Lent, L., Vilkov, G., & Zhang, R. (2023). Firm-level climate change exposure. *The Journal of Finance*, *78*(3), 1449–1498 (cit. on p. 144).

Schoenfeld, J. (2017). The effect of voluntary disclosure on stock liquidity: New evidence from index funds. *Journal of Accounting and Economics*, *63*(1), 51–74 (cit. on p. 133).

Scott, T. (1994). Incentives and disincentives for financial disclosure: Voluntary disclosure of defined benefit pension plan information by canadian firms. *The Accounting Review* (cit. on p. 147).

Shroff, N., Sun, A. X., White, H. D., & Zhang, W. (2013). Voluntary disclosure and information asymmetry: Evidence from the 2005 securities offering reform. *Journal of Accounting Research*, *51*(5), 1299–1345 (cit. on p. 132).

Siew, R., Balatbat, M., & Carmichael, D. (2016). The impact of esg disclosures and institutional ownership on market information asymmetry. *Asia-Pacific Journal* *of Accounting and Economics*, *23*(4), 432–448. https://doi.org/10.1080/16081625.2016.1170100 (cit. on p. 142).

Stoll, H. R. (2000). Presidential address: Friction. *The Journal of Finance*, *55*(4), 1479–1514 (cit. on p. 148).

WBCSD. (2019). Reporting matters 2019 [Accessed: 2025-01-15]. *World Business Council for Sustainable Development*. https://docs.wbcsd.org/2019/10/WBCSD_Reporting_Matters_2019.pdf (cit. on p. 134).

Wong, W., Batten, J., Mohamed-Arshad, S., Nordin, S., & Adzis, A. (2021). Does esg certification add firm value? *Finance Research Letters*, *39*, 101593 (cit. on p. 137).

Zhang, L., Shan, Y., & Chang, M. (2021). Can csr disclosure protect firm reputation during financial restatements? *Journal of Business Ethics*, *173*, 157–184 (cit. on p. 132).

Chapter 4

Democratizing sustainability assessment: Leveraging generative AI for assessing corporate sustainability

Melanie Luxembourger* / Maxime Pettinger†

University of Luxembourg /STATEC, Luxembourg

Abstract

This study explores the potential of leveraging generative artificial intelligence (AI), specifically ChatGPT, for assessing corporate sustainability. We examine ChatGPT's reliability in sustainability scoring and investigate the factors contributing to discrepancies between AI-generated and a composite benchmark from established traditional data providers, namely Refinitiv, Bloomberg, and S&P500.Our methodology includes automating ChatGPT to evaluate companies based on emissions, innovation, and resource use, comparing these scores to a composite of the three benchmark scores (RBS), and assessing the associated misclassification risk. We also explore whether ChatGPT's suitability varies based on firm-specific characteristics such as size, geographic location, industry sector, age, and profitability. Our findings reveal that ChatGPT can generate sustainability assessments, but with limitations. The AI shows a high non-response rate and significant differences in score distributions compared to RBS. ChatGPT tends to produce higher, more concentrated scores and struggles with firms exhibiting lower environmental performance. Despite these challenges, there exist an association between ChatGPT and RBS, indicating partial overlap in informational content. Certain firm characteristics, including larger size, European location, and older age, enhance the reliability of ChatGPT's assessments. These results emphasize the need for more openly available and regulated sustainability data, as well as the importance of continued innovation and improvement in AI technology to enhance the coverage and accuracy of such assessments. This research contributes to a new approach to assessing corporate sustainability, where advanced technologies and open, unbiased data support businesses in their sustainability journey.

Keywords— Corporate Sustainability, ESG Ratings, Generative AI, Sustainability assessment

1 Introduction

Corporate sustainability has become a cornerstone of modern business practice, driving demand for reliable tools to assess and improve sustainability performance. Traditionally, businesses have relied on data providers and rating agencies to provide the necessary metrics to gauge environmental, social, and governance (ESG) achievements. However, these services often require long-term subscriptions and represent significant costs, making them inaccessible to many small and medium-sized enterprises (SMEs) and investors. This research paper explores leveraging generative AI to mitigate these challenges. Specifically, we consider using ChatGPT, a generative AI model, to evaluate corporate sustainability. This approach offers a cost-effective, customizable, and flexible solution, simulating human-like interaction through its chat interface, making it accessible to a broader range of stakeholders. This research investigates whether ChatGPT is reliable for sustainability scoring and whether it could replace traditional ESG rating agencies. Additionally, we examine the factors explaining discrepancies between AI-generated and traditional scores, and identify specific groups of firms for which ChatGPT is particularly effective. Our methodology involves retrieving environmental scores from three data providers—Refinitiv, Bloomberg, and S&P500. We then automate ChatGPT to evaluate companies against the same criteria—focusing on emissions, innovation, and resource use. We compare ChatGPT's environmental scores with a composite of the three benchmark scores (RBS) and characterize the mapping between the two, including the use of various association measures. We then assess the misclassification risk associated with using ChatGPT for environmental performance assessment. Finally, we evaluate the determinants of discrepancies between RBS and ChatGPT scores, exploring whether ChatGPT's suitability is context-dependent. Given that ChatGPT's ability to provide an assessment and its accuracy is influenced by the amount of information it has access to, and because signaling theory suggests that firms with high sustainability performance tend to communicate more about their practices, we investigate how firm-specific characteristics that impact sustainability performance (e.g., size, geographic location, industry sector, age, and profitability) affect AI's ability to provide an assessment and the accuracy of such assessments. To identify the determinants of discrepancies, we rely on generalized linear models. Our investigation reveals that ChatGPT can evaluate sustainability performance, but only for a small fraction of the companies for which it was requested. The AI-generated scores are only modestly associated with traditional environmental scores, despite the expectation that the two measures should represent the same underlying concept. We also find that ChatGPT's evaluations are more reliable for firms exhibiting specific characteristics, such as larger size, European location, older age, and certain industry sectors. These findings highlight both the potential and limitations of using generative AI for sustainability assessments. ChatGPT can aggregate and analyze information effectively, but its success relies on the quality and accessibility of the data it processes, emphasizing the need for more openly available sustainability data. Therefore, third-party evaluators remain essential to ensure comprehensive and accurate assessments. This research aims to contribute to a new approach to corporate sustainability, where advanced technologies and open, unbiased data work together to support businesses in their sustainability journey.

2 Prior literature

Companies disclose their data through various means, such as sustainability reports, websites, and brochures, reflecting the evolving nature of corporate communication on sustainability and growing stakeholder interest (Michelon et al., 2015). Despite advancements in corporate transparency, the usability and accessibility of sustainability data remain challenging. Data providers (e.g. MSCI, Sustainalytics, Bloomberg, Refinitiv ISS ESG, etc.), while enabling the incorporation of non-financial metrics into corporate evaluations, use different methodologies involving proprietary measurements, scope, and weighting of ESG factors, and diverse data sources, ranging from public information to proprietary research and third-party data, leading to discrepancies in ESG scores and complicating comparisons due to each agency's unique approach (Berg et al., 2022; Billio et al., 2021; Chatterji et al., 2016; Christensen et al., 2022; Dimson et al., 2020). Additionally, these methodologies often overlook qualitative and non-financial information, hindering comprehensive sustainability evaluations (Kotsantonis and Serafeim, 2019). Moreover, while essential for informed decision-making, their high costs restrict access for smaller entities, widening the gap between large corporations and SMEs (Hoogendoorn et al., 2019).

One potential solution to these challenges is the concept of open data, defined as information provided in a machine-readable digital format that is easily accessible, free, and open to the public (Helbig et al., 2021). However, merely making data open is insufficient. Stakeholders may be overwhelmed by the sheer volume of data available, as they have limited time and resources to analyze large datasets. Additionally, the aggregation and dimensionality reduction necessary to transform raw data into actionable insights present significant challenges, as traditional data providers demonstrate. This process is often costly due to the extensive human effort involved.

To address these challenges, integrating AI offers promising solutions. In the last 30 years, research on AI in accounting has steadily increased (Sutton et al., 2016), as the field faces numerous complex tasks that can benefit from AI, such as automating physical tasks, analyzing numbers through algebraic analysis, visual analytics, and hypothesis-based predictive analytics (Kokina and Davenport, 2017). Generative AI, like ChatGPT, shows significant potential in simplifying complex disclosures and improving interpretability across various fields (Ayinde et al., 2023),helping stakeholders enhancing decision-making by simplifying financial disclosures, reducing document length, and emphasizing key information (Kim et al., 2024). Building on these capabilities, researchers have explored AI's role in automating and improving corporate sustainability analyses. Large language models (LLMs) like the one underlying ChatGPT can process textual information from sustainability reports, extract insights, and enhance knowledge graphs for better representation of sustainability data (Ni et al., 2023). With their extensive training sets encompassing diverse textual data, these models transform unstructured data into structured formats to streamline information processing and enhance the effectiveness of sustainability assessments. Practical applications involve using firm-specific inputs, such as financial disclosures or sustainability reports, to tailor outputs to particular contexts (Ayinde et al., 2023; Kim et al., 2024; Ni et al., 2023). This ensures generated analyses and summaries are comprehensive and relevant by combining the model's pre-trained capabilities with firm-specific data. While most research focuses on providing ChatGPT with structured datasets, there is a need to explore its autonomous capabilities in real-world scenarios where specific data might not always be readily available. In this context, our research addresses two main issues: the high cost of traditional ESG evaluation services and the accessibility of the evaluation process. By leveraging ChatGPT's ability to aggregate and process large volumes of publicly available data, we consider a cost-effective and user-friendly alternative for sustainability evaluations. This approach not only makes sustainability data more affordable and accessible but also democratizes sustainability assessments, empowering a broader range of stakeholders, especially those lacking the resources for traditional, high-cost aggregation services.

3 Hypothesis development and strategy

In this study, we explore the viability of using ChatGPT for sustainability scoring by comparing its performance to a composite score derived from three well-established ESG evaluation services: Refinitiv, Bloomberg, and S&P500 (RBS hereafter). Our investigation is guided by two primary research questions: (Q1) Is ChatGPT reliable for ESG scoring? and (Q2) What factors contribute to the discrepancies, if any, between ChatGPT's scores and those of traditional ESG rating agencies? To address these questions, we propose hypotheses to systematically evaluate ChatGPT's performance in ESG scoring.

3.1 Scoring capacity

Addressing Q1 involves examining the similarity between firm ratings given by ChatGPT and RBS. It requires examining and quantifying the association between the scores. We proceed in two steps: First, we compare the source-of-scoring-specific univariate distributions to determine if the two statistical objects are similar. Due to inherent variations in the scoring systems' mechanisms, we anticipate disparities in the variables produced. ESG rating agencies use advanced dimensionality reduction techniques to linearly combine many ESG-related factors into a single composite index, on a continuum from 0 to 100 for Refinitiv and S&P500 and 0 to 10 for Bloomberg. Conversely, ChatGPT uses its text prediction abilities to produce new text tailored to the user's given context. Following these characteristics, we propose the first hypothesis:

H1: Given ChatGPT's design to process and generate human-like text, its scoring scale is akin to a Likert scale ranging from 'very poor' to 'excellent'.

Second, we assess the degree of association between the two metrics and the overall risk of misclassification implied by using ChatGPT. We employ various composite scores and measure of association, and examine the mapping between the two scoring scales, focusing on different segments of the support, or 'categories'.

3.2 Factors explaining discrepancies

Addressing Q2 involves examining the impacts of various factors on both the probability of ChatGPT generating a score and the distance from RBS scores. Analyzing these factors helps identify categories of firms where ChatGPT performs best or worst.

3.2.1 Ability of ChatGPT to generate a relevant score

To generate a score for a specific firm, ChatGPT must be trained on a corpus of texts where this firm has been (repetitively) associated with an assessment of its ESG practices, especially if prompted to avoid reliance on general information. Larger datasets significantly enhance the performance of the model, so much that smaller models with bigger training sets have been able to outperform larger models with smaller training sets (Hoffmann et al., 2022). However, the quality of data is critical, as corporate disclosures are often designed to highlight positive performance and downplay negative aspects

(MerklDavies et al., 2011), and firms with significant media influence can shape public perception through selective disclosure and positive framing (Christensen et al., 2022), resulting in biased data in AI models. This is further reinforced by the model's design, which aims to produce socially acceptable outputs (Binns et al., 2018). Nevertheless, regulations and enforcement mechanisms can lead to better-quality disclosures and improved information environments, mitigating those challenges (Krueger et al., 2021). Therefore, ChatGPT should only generate a score if its training set includes a significant amount of information about this firm, and the score would be reliable is the data it had access to was of quality. In their current form, these statements are hardly empirically testable. Building a direct firm-specific measure of the quantity and quality of publicly available ESG information would require access to the same corpus of texts used to train ChatGPT, which is beyond the scope of most organizations. However, both the quantity and the quality of information available should depend on the visibility of the firm and the resources it allocates to publishing such information. To proxy for firm visibility and resource allocation, signaling theory suggests that firms with superior environmental performance are more likely to disclose their achievements comprehensively to distinguish themselves from competitors (Clarkson et al., 2008; Dhaliwal et al., 2011; Eccles et al., 2014). These disclosures signal a firm's commitment to sustainability, attracting investment and enhancing its reputation. Consequently, firms with strong ESG performance are more likely to publish comprehensive information about their initiatives.

3.2.2 Determinants of the quantity and quality of publicly available ESG information

To identify relevant variables that explain a firm's ESG performance, we draw on the literature and identify the following factors:

Firm Age: Mature firms tend to invest more in ESG (Diebecker et al., 2017), making them more likely to have information available about their sustainability strategies.

Firm Size: Larger firms, measured as the number of employees (Aguinis and Glavas, 2012), revenues (Morhardt, 2010), market value (Jennifer Ho and Taylor, 2007), or market capitalization (Garcia et al., 2019), face increased pressure from public exposure and stakeholder expectations to act responsibly and sustainably, driven by their greater resources and greater accountability.

Geographic Location: European companies face high expectations from the public and stakeholders, with a cultural emphasis on sustainability and social responsibility (Gjølberg, 2009; Matten and Moon, 2008), which fosters superior performance and stronger regulatory frameworks, resulting in higher-quality sustainability reports (Dilling, 2010).

Financial Performance: Research indicates a mixed relationship between profitability (ROA) and ESG performance and disclosure (Drobetz et al., 2014), with some studies finding a positive relationship due to enhanced resources (Diebecker et al., 2017; Marquis and Qian, 2014), while others suggest that less profitable firms might increase ESG disclosure to offset poor financial performance (Jennifer Ho and Taylor, 2007).

Industry: High-profile sectors like energy and production face greater scrutiny and pressure to demonstrate sustainability, leading to high-quality, detailed sustainability reports to meet stakeholder expectations and maintain their market position (Dilling, 2010).

For each of these variables, we evaluate the existence and direction of a non-negligeable relationship with environmental performance within our sample by including them as explanatory variables in regressions models (pooled OLS and a panel linear model (PLM) with individual fixed effects).

We then use these variables in a logit model predicting the probability of GPT generating an answer; and form the following hypothesis:

H2: There is a significant positive relationship between variables considered as linked to the quantity of published information and the probability of ChatGPT generating an answer.

Concerning the precision of ChatGPT scores, which is inversely proportional to their deviation from RBS scores, we employ a linear regression model to predict the distance between ChatGPT and RBS scores and form the following hypothesis:

H3: The distance between ChatGPT and RBS scores has a significant negative relationship with the variables that we consider as linked to the quantity of quality of publication of information (i.e., they reduce the distance), and a positive significant relationship with ChatGPT's non-response rate.

Finally, we expect that factors impacting a firm's ESG performance also significantly impact its propensity to make publicly available (textual) information about such performance. These factors should similarly impact both ChatGPT's propensity to provide an answer and the precision of this answer, leading to the following hypothesis:

H4: Variables that have a significant impact on ChatGPT's scoring performance (H2) and precision (H3) are consistent in sign with those observed as having a non-negligeable relationship with environmental performance.

4 Sample

4.1 Sample selection

To establish our study's dataset, we began by retrieving the Environmental Pillar Scores from Refinitiv's comprehensive ESG database (LAST4ESG series), which covers over 90% of the global market cap across more than 630 metrics with historical data dating back to 2002. We focused on firms with ESG scores available in 2022, resulting in a dataset of 10,353 companies worldwide. This selection ensures our analysis captures both historical trends and contemporary practices[1].

4.2 ChatGPT's evaluation collection

The user prompt was sent using the *openai.ChatCompletion.create*[2] method, and the script extracted the company name, numerical sustainability score, and detailed assessment using regular expressions, compiling the data into a dataset.

[1]One limitation to note is that the sample includes only companies active in 2022, introducing survivorship bias by excluding firms that ceased operations earlier, which may have had differing sustainability practices or ESG strategies.

[2]Please refer to Appendix A for more detail.

However, ChatGPT's underlying stochastic architecture implies that it might generate different outputs for the same input across different runs. To address this variability, we conducted ten simulations for each company, allowing us to observe the range of potential outputs and determine the assessments' consistency. Despite these efforts, the inherent randomness remains a limitation that can impact the reproducibility of the results.

4.3 Composite index creation

For each firm with a 2022 Refinitiv score, we then gathered corresponding environmental scores from two additional data providers: Bloomberg Environmental Pillar Score and S&P Global's Environmental Dimension Rank. One challenge posed here relies on the availability of the scores. Indeed, not all providers started providing scores at the same time, and they do not cover all the same firms.

[Insert Table 4.1 about here]

Our methodology to create a composite score comparable with ChatGPT involved two key steps: combination of agency scores at the firm-year level and period-level aggregation.

4.3.1 Firm-year combination

In the first step, composite scores are created for each firm on a yearly basis, combining scores from the three agencies: Refinitiv, Bloomberg, and S&P500. To account for data limitations and maximize the utility of available information, we also constructed subsets of composite scores, namely RB (Refinitiv and Bloomberg), RS (Refinitiv and S&P500), and BS (Bloomberg and S&P500). For each score combination, two approaches were applied: an **opportunistic method**, which includes all available scores per given firm-year combination, and a **pure method**, which requires that all agencies considered provide scores for a firm each year. If any score is missing, the composite score for that firm-year is not calculated. While this ensures a consistent definition of the composite score, it reduces the sample size, particularly in earlier years when data coverage was limited.

Under each approach, three methods were used:

- **Simple average**, where the scores are equally weighted.
- **PCA-based weights (pooled)**, where principal component analysis (PCA) is applied across all years (i.e. "pooled"), to determine the weights for for each agency's score.
- **Year-specific PCA weights**, where PCA is applied separately for each year, which has the advantage of capturing the evolution in common variance across time, due for instance to changes in agency-specific methodologies.

4.3.2 Period-level aggregation

One key limitation of the ChatGPT model used in this study is its inability to provide year-specific scores. This arises from the model's architecture, which generates responses based on patterns in aggregated data, rather than on specific temporal information. To address this limitation, we create a global composite score averaging over a given period. This approach involves a critical trade-off: longer periods are more likely to mimic ChatGPT's training data, making the metrics more comparable; however, averaging across different years hide dynamic changes in a firm's policies and environmental scores over time.

Besides the firm-year combination discussed previously, which preserves the panel aspect of the original data, we incorporate two additional methods for firm-level analysis. The first method simply takes the average, for each firm and over-the-period, of the firm-year combination. This method is also applied to the individual firm-year scores provided by each agency, to provide baseline comparisons and to evaluate the standalone performance of each agency. The second method takes the average for each firm and over the period for each agency, and then combine using PCA to determine the weights for combining these averages into the final score. Table 2 summarizes all the aggregation methods considered for the composite score.

[Insert Table 4.2 about here]

For each aggregation method outlined in Table 2, we calculate the composite scores using various lower bounds for the starting year, beginning from 2002 onwards. This equals to a total of 770 benchmark scores considered for the analysis:

$$\begin{aligned}\big[&(8 \text{ possible aggregation methods} \times 4 \text{ possible score combinations}) \\ &+3 \text{ single agency scores}\big] \\ &\times 22 \text{ possible year starts} \\ &= 770 \text{ benchmark scores.}\end{aligned}$$

4.3.3 Choosing a baseline model and period

Our primary aim is to maximize the availability of data while ensuring a representative composite score. As such we first select RBS for our score combination, since a multi-source composite provides a more holistic and balanced assessment of environmental performance as it mitigates the influence of methodological biases or discrepancies inherent in any individual rating system. Then, the opportunistic approach emerges as a practical choice, as it leverages all available scores, even when data from one or more agencies is missing, minimizing data loss and, by extension, increasing statistical power. Figure 1 illustrates how the different approaches impact the informational content of the methods considered. The figure shows that the opportunistic approach captures proportionally more variance from Refinitiv, particularly in earlier years when Bloomberg and S&P data are sparse. This highlights the trade-off between maximizing data availability and maintaining uniformity in the composite's informational content.

[Insert Figure 4.1 about here]

Within the opportunistic framework, EAPY is selected as our baseline model for its dynamic adjustment to year-specific variations in data availability, while allowing the composite score to adapt to changes in the common variance between agency scores and ensuring that it reflects temporal shifts in the relative contributions of individual agencies while preserving as much data as possible. While this approach may give disproportionate weight to Refinitiv in earlier years—owing to its stronger presence during those periods—this is a calculated trade-off as it maximizes the number of observations, enhancing the reliability of the statistical models used in the analysis, particularly for longitudinal assessments.

Concerning the starting year for aggregation, while some arbitrariness in selecting it is unavoidable, we developed a decision criterion to enhance transparency.

[Insert Figure 4.2 about here]

Figure 2 shows the distribution of EAPY_RBS scores over different periods and aggregates (by firm and overall). Over a shorter period (e.g., 2020 to 2023), data dispersion is less significant compared to a longer period. Longer periods capture significant trends and changes in corporate strategy and sustainability practices, essential for understanding ESG performance variability. Therefore, longer periods theoretically align better with ChatGPT's aggregated data approach, enhancing model relevance and comparability.

[Insert Figure 4.3 about here]

Moreover, we find in figure 3 that the most significant reduction in firm-specific dispersion occurs when starting from 2016. This year also marks one of the first periods where all three agencies provided consistent coverage, ensuring a more robust dataset. Therefore, we chose 2016 as the starting year for our baseline, serving at the baseline period (2016-2023).

5 Results

The primary objective of leveraging ChatGPT in our study is to ascertain its capability to autonomously assess corporate sustainability and to gauge the accuracy of these assessments against established ESG ratings providers. This sub-section presents the findings of ChatGPT's performance as an assessment tool in terms of capacity and alignment with traditional methods.

5.1 Capacity of ChatGPT to provide an assessment

GPT's ability to generate sustainability assessments is assessed by configuring it to analyze sustainability aspects such as emissions, resource usage, and environmental innovation. Initially, using its training data, the model produced some "hallucinations" due to a lack of specific information about the entity (i.e. "the evaluation provided here is based on general practices and trends within the industry, as well as the broader context of "Company"'s operations, to infer potential areas of focus and improvement."). These hallucinations, i.e. plausible but incorrect information (Maynez et al., 2020), were deemed unreliable. As such the prompt was modified to be more restrictive and limit the hallucinations (please refer to Appendix A for more detail).

[Insert Table 4.3 about here]

Analysis of the responses under both original and revised prompts in table 3 showed a high rate of non-responses, indicating a significant limitation in ChatGPT's capacity to autonomously generate relevant sustainability assessments without detailed, specific input. This is consistent regardless of the prompt's restrictiveness, suggesting challenges related to the complexity of the task and the model's reliance on the data available within its training set.

5.2 Comparison of univariate distributions

To evaluate the reliability of ChatGPT's sustainability scores, we conducted a comparative descriptive analysis to determine if the distributions of scores for ChatGPT and RBS scores exhibit similar statistical characteristics. This analysis characterizes the univariate distribution of the two scores, examining their alignment in terms of central tendency, variability, and overall data patterns.

[Insert Figure 4.4 about here]

Figure 4 shows significant differences between RBS and ChatGPT's score distributions in terms of overall shape, position, and dispersion. RBS scores have a broader, more skewed distribution, spanning a continuous range from 0 to 100, with a mode at zero, indicating many firms with low ESG scores. In contrast, ChatGPT-generated scores are higher, more concentrated, and seemingly multimodal, suggesting these scores might be treated as ordered multinomial variables.

This suggests the model may categorize assessments into discrete categories, potentially reflecting a Likert scale-like type of scoring system, as proposed in H1. The clustering of scores may result from ChatGPT being a text predictor rather than a precise data analyzer. ESG rating agencies like the ones considered in RBS combine many dimensions into one composite indicator (see 3.1). However, ChatGPT likely attempts to identify a scale measure in the text, quantified as 'disastrous,' 'bad,' 'neutral,' 'good,' 'excellent,' aligning with the multimodal peaks in the kernel density estimate. This behavior is more evident in individual runs; however, computing the average scores (Figure B) mitigates this effect, smoothing the distribution across the entire range. Figure D shows that when considering only RBS scores for observations where ChatGPT generated at least one score, the distribution shifts, indicating ChatGPT's tendency to omit scores for firms with low ESG performance. ChatGPT may struggle to evaluate companies with lower scores due to insufficient data, as lower-scoring firms tend to report less about their practices, while higher-scoring firms are more likely to publicly promote their sustainability efforts. Moreover, the apparent positivity in ChatGPT's scores might be influenced by non-response bias where ChatGPT assigns not-available (NA) instead of lower scores, inflating the estimated metrics of position.

5.3 Association between ChatGPT and RBS scores

As Figure 4 suggests the existence of five density peaks, consistent with a traditional Likert scale, we recode the mean GPT-generated scores as ordered categorical variables, using the density peaks as centers of the ranges defining the categories. However, ChatGPT tends to fail in generating scores for firms with poor ESG performance, suggesting that the absence of a score itself provides information about a firm's ESG performance. Thus, we create a version of the categorically recoded ChatGPT score that considers NAs as a sixth, lowest category.

5.3.1 Overall assessment of association

[Insert Figure 4.5 about here]

The analysis in Figure 5 highlights a modest relationship between ChatGPT and benchmark scores, with Pearson's r(median value: 0.356) and Spearman's ρ (median value: 0.335) showing weak correlations when both scores are treated as continuous variables. However, our previous analysis revealed that ChatGPT scores are more akin to ordinal variables than continuous metrics. To explore this, other boxplots treat ChatGPT's scores as ordered multinomial variables. When analyzed using Spearman's ρ, the correlation remains modest (0.328–0.329) for both 5-category and 6-category ordinal recordings of ChatGPT's scores. What significantly increases the correlation is the inclusion of a missing score category ("NA category"), achieving the highest point estimate of 0.408. This underscores the informational value embedded in ChatGPT's inability to generate a score, indicating that non-responses often correlate with poor ESG performance. This suggests that treating missing ChatGPT's scores as a distinct category provides meaningful insights into the firm's ESG standing. While ChatGPT's scoring does not align perfectly with the benchmark scores, these results indicate the presence of shared informational content between the two systems.

[Insert Table 4.4 about here]

Moreover, table 4 highlights that the correlation coefficients between the actual data providers considered in the study are themselves not very strong, showing correlation coefficients ranging from 0.449, a moderate value, up to 0.624, which approaches a strong correlation, but remains close to moderate in strength. This contextualizes the modest correlations observed between ChatGPT and the benchmark scores. Under certain conditions, ChatGPT could potentially serve as a substitute for an ESG-rating agency, at least for specific segments of the ESG performance spectrum. The following section provides a more detailed assessment of the alignment between ChatGPT and RBS scoring, including an indirect quantification of the misclassification risk associated with relying on ChatGPT for environmental performance assessment.

5.3.2 Misclassification risk analysis

To illustrate ChatGPT's deviation from optimal scoring, we compare ChatGPT's categorical scoring with an ideal interval scale. Interval measures not only have the properties of distinctiveness and order—as ordinal measures do—but also have the additional property of equal intervals property of equal intervals (Bandalos, 2018). This makes interval scales more closely aligned with continuous measures, which is the type of scoring system that RBS provides. This comparison highlights the misclassification risk by benchmarking ChatGPT against a more precise system. We divided the continuous range of scores into six equal-sized intervals, each representing a distinct category. This ensures a perfect mapping between the created categories and the RBS scores, enhancing the accuracy of our comparison.

[Insert Figure 4.6 about here]

Figure 6 shows that GPT categories have heterogeneous sizes and significantly overlap with the distribution of RBS scores within each category. For instance, categories two, three, and four have similar mean and median RBS scores and wide within-category dispersion. Each category spans a broad range of RBS scores, indicating a high risk of misclassification. Category one is particularly notable for its broad and uneven distribution of RBS scores, suggesting that ChatGPT combined observations from different RBS ranges into this intermediate category, likely including firms from both category zero and category two.

[Insert Figure 4.7 about here]

Figure 7's left panel provides a 'success' rate for each possible degree of misclassification. A degree of zero implies that ChatGPT has categorized the firm in the same category as the optimal scoring system. This is the case for 35.78% of the firms, indicating a raw misclassification rate of 64.22%. While this misclassification rate remains significant, it is notably lower than that of a purely random scoring system ($100 \times (1 - 1/6) = 83.33\%$, six being the total number of categories). Other bars indicate 'failure,' where firms are classified by ChatGPT into different categories than the optimal ones. Moreover, unlike random scoring, the frequency of deviations decreases with their intensity, reflecting the ordered nature of ChatGPT categories.

The right panel's augmented confusion matrix provides a deeper analysis of specific misclassifications between ChatGPT and Optimal categories, addressing the earlier contradiction where ChatGPT's underestimation seemed at odds with findings in Section 5.2, Figure 4. Notably, firms in category zero are over-represented, reflecting cases where ChatGPT failed to generate a score. As the lowest category, this indicates either correct classification or underestimation, explaining the positive differences seen in the left panel and partially resolving the contradiction with the earlier findings.

5.4 Factors impacting firms' ESG disclosures existence and precision

5.4.1 Investigating discrepancies

Even after excluding not-available (NA) entries, both the mean and median differences between ChatGPT and RBS scores remain positive, indicating a systematic deviation where ChatGPT tends to assign higher scores. The distribution of the differences between ChatGPT and RBS scores is positively skewed, further underscoring this tendency. The histogram in Figure 8 shows that most differences are concentrated above zero, reflecting this positive bias.

[Insert Figure 4.8 about here]

5.4.2 Analysis of determinants of ESG Scores

[Insert Figure 4.9 about here]

Building upon the potential factors identified in the literature as contributing to discrepancies, this section seeks to validate these factors and estimate their directional impact on environmental scores. Specifically, figure 9 examines the influence of various firm- and industry-level determinantsPlease refer to Appendix B for the definition of the variables. on the benchmark scores which capture the environmental performance of a firm. Consistent with previous research, the results demonstrate that the following variables—being European (`eur`), older (`age`), several measures of size such as `ln_employees`, `ln_market_value` and `ln_revenues`, and being in the real estate sector (`industryRealEstate`)—consistently have positive and significant coefficients with environmental scores. Additionally, ChatGPT's non-response rate (`na_rate_gpt4c`) is strongly negatively correlated with environmental performance, demonstrating the relationship between lower environmental performance and decreased ChatGPT's likelihood of generating a response, thereby partially confirming H4.

5.4.3 Comparative analysis of ChatGPT and RBS scores discrepancies

Our subsequent analysis focuses on evaluating the effectiveness of the identified variables in explaining the discrepancies between ChatGPT and benchmark scoresIn cases where variables or weights are derived from the data itself, such as using PCA to compute weights, there is inherent additional variability. This variability can compromise the assumptions necessary for standard statistical inference methods, which rely on assumptions like data being independent and identically distributed (i.i.d). Since these assumptions might be violated, bootstrapping, which resamples the data to create many simulated samples, can provide more robust and reliable estimates. As such, to enhance the robustness of our findings, the coefficient estimates graphs with bootstrapped standard errors are presented in Appendix C..

[Insert Figure 4.10 about here]

[Insert Figure 4.11 about here]

In our analysis of Figure 10 and 11, size, the age, location, and industry consistently influenced the performance and reliability of ChatGPT scoring across all models. Larger companies (i.e., with more employees), older companies, European companies, and those in the real estate sector are more likely to have a ChatGPT score, with positive significant coefficients, in line with H2. These variables are also associated with greater precision, as evidenced by smaller differences between ChatGPT and RBS scores and reduced overestimation intensity, thereby aligning with H3.

However, discrepancies arise in Model 3, where only eur remains significant within the EAPY_RBS_2016 benchmark selected for the study. Nonetheless, `age`, `ln_employees`, and `industryRealEstate` maintained the expected negative impact on overestimation across all benchmarks, except for a few observations concerning `ln_employees`. These variables are also significant in a substantial percentage of benchmarks (26%, 38.8% and, 28.2% respectively). The smaller sample size in Model 3 (average of 722 observations compared to 7,396 for Model 1) likely restrict statistical power, reducing the estimates' precision.

As such, these variables that have a significant relationship with environmental performance also enhance ChatGPT's ability to provide relevant and accurate scores. This observation is consistent with H4, which posits that factors impacting a firm's ESG performance also affect the availability and quality of publicly accessible ESG information, ultimately shaping ChatGPT's scoring capabilities.

Our analysis also reveals that ChatGPT's non-response rate (`na_rate_gpt4c`) significantly influences the differences in scores and overestimation likelihood and severity, with a positive coefficient across all models, indicating a trend of greater discrepancies with higher NA rates. A higher ChatGPT NA rate suggests the model struggled to find relevant information, often producing outputs that were not directly pertinent to the query, leading to overestimation due to assumptions or generalizations rather than specific, accurate data (Bender et al., 2021). This phenomenon, known as "hallucination," involves generating plausible but incorrect information (Maynez et al., 2020).

6 Concluding remarks

This study investigated the potential of leveraging generative AI, specifically ChatGPT, for assessing corporate sustainability. We examined the reliability of ChatGPT for ESG scoring and the factors contributing to discrepancies between AI-generated and traditional scores. The study revealed a high non-response rate in ChatGPT's ability to generate sustainability assessments, indicating challenges due to task complexity and data dependency. Moreover, there are significant differences between ChatGPT and RBS scores in their univariate distributions. Traditional data providers' scores have a broader distribution, with a high representation of low ESG firms. GPT-generated scores are usually higher, more concentrated, and exhibit a multimodal distributional pattern, possibly reflecting a Likert scale scoring system. ChatGPT often omits scores for firms with low ESG performance, indicating a struggle to assess companies with lower scores. Investigating the association between RBS and ChatGPT scores, we found a modest relationship between the two, with a substantial risk of misclassification. The existence of a non-negligible association between the two scoring systems nonetheless indicates that they partially share informational content. Therefore, ChatGPT could possibly be used as a substitute for an ESG-rating agency, at least for some segments of ESG performance. Our study indeed shows that firm characteristics such as size, age, geographic location, and industry significantly influence both the likelihood of ChatGPT generating a score and the accuracy of these scores compared to RBS.

These findings highlight both the potential and limitations of using generative AI for sustainability assessments. While ChatGPT shows promise as a cost-effective and customizable solution, it is not yet a comprehensive substitute for traditional ESG rating agencies. As a generalist LLM model, ChatGPT has inherent limitations in assessing corporate sustainability: it is not specifically designed for building composite indicators and is dependent on the availability and quality of data included in its training set. As such, the model lacks precision for specific assessments and produces limited output, particularly for firms with less publicly available information.

However, as models improve and potentially specialize, and as regulatory frameworks evolve (i.e., more high-quality data becomes available), generative AI is likely to become more relevant for this type of exercise. Moreover, future research could focus on developing methodologies to mitigate biases in AI training data.

7 Figures and Tables

Table 4.1: Data availability for environmental scores, including number of observations and completion rates, by year and agency.

Year	Bloomberg	Refinitiv	S&P500
2002	0 (0%)	525 (5%)	0 (0%)
2003	0 (0%)	531 (5%)	0 (0%)
2004	0 (0%)	1028 (10%)	0 (0%)
2005	0 (0%)	1318 (13%)	0 (0%)
2006	0 (0%)	1342 (13%)	0 (0%)
2007	0 (0%)	1510 (15%)	0 (0%)
2008	0 (0%)	1821 (18%)	0 (0%)
2009	0 (0%)	2104 (20%)	0 (0%)
2010	0 (0%)	2541 (25%)	0 (0%)
2011	17 (0%)	2695 (26%)	0 (0%)
2012	21 (0%)	2805 (27%)	2 (0%)
2013	38 (0%)	2924 (28%)	5 (0%)
2014	116 (1%)	3064 (30%)	10 (0%)
2015	3811 (37%)	3675 (36%)	33 (0%)
2016	3918 (38%)	4300 (42%)	1041 (10%)
2017	4058 (39%)	5248 (51%)	2408 (23%)
2018	4169 (40%)	5985 (58%)	3303 (32%)
2019	4264 (41%)	7034 (68%)	4129 (40%)
2020	4504 (44%)	8270 (80%)	4444 (43%)
2021	8535 (81%)	9340 (90%)	6452 (62%)
2022	8351 (80%)	10347 (100%)	6444 (64%)
2023	7980 (77%)	7673 (74%)	6816 (66%)

Table 1 summarizes the availability of environmental scores per year, detailing the number of observations (i.e., firms with environmental scores) and the completion rates (%) for each agency. The data availability varies considerably across years and providers. Before 2016, we have virtually no environmental score from Bloomberg and S&P500 for the firms included in our sample.

Figure 4.1: Correlation between the composite and agencies indicators for each starting year

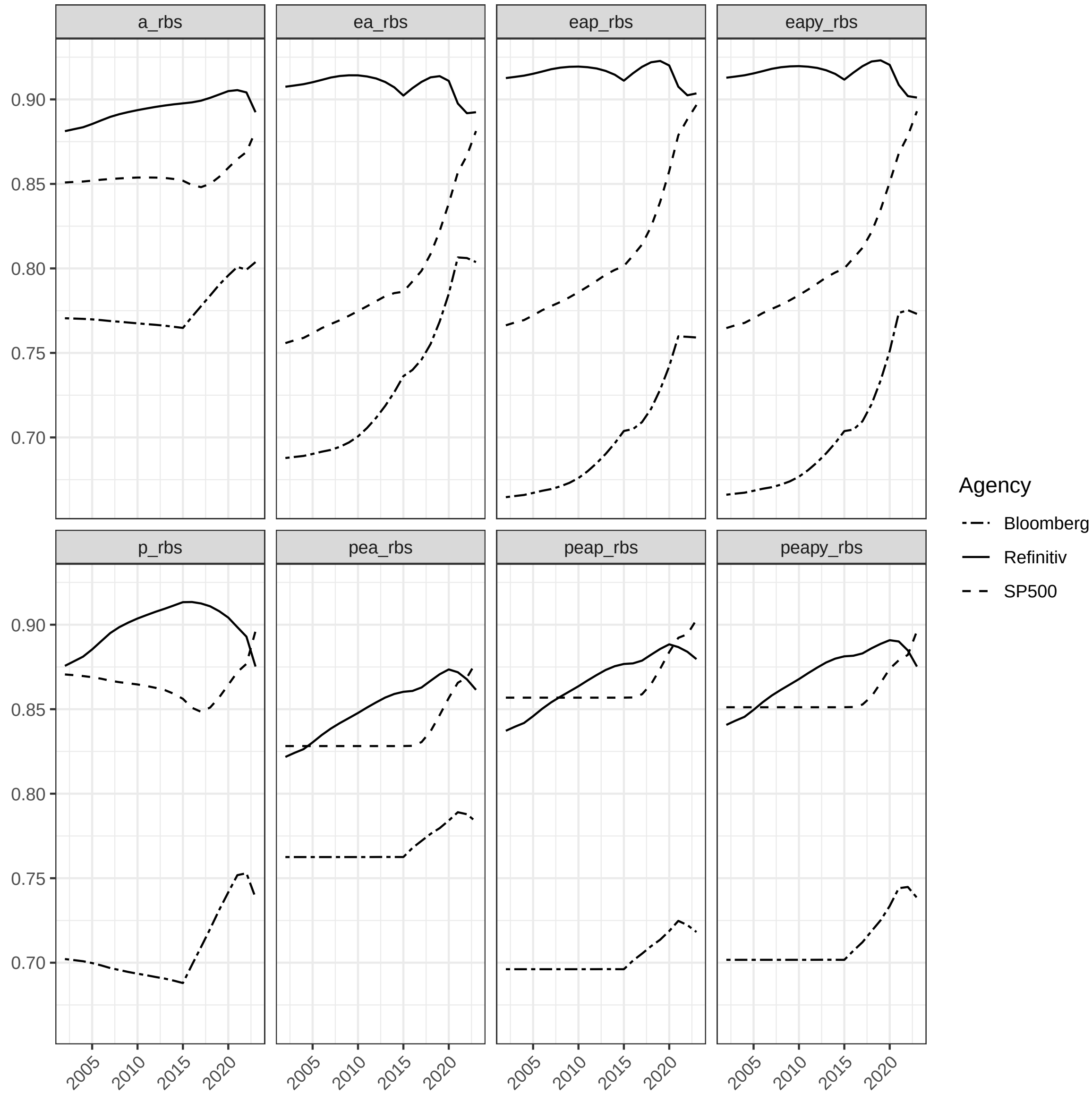


Figure 1 illustrates the correlation between composite indicator RBS and agency-specific metrics over various starting years. It highlights the differing impacts of opportunistic versus pure strategies on the over-period average. The results demonstrate that the opportunistic approach consistently yields higher correlations with Refinitiv scores, likely due to its flexibility in accommodating missing data by relying on available sources, which often prioritize Refinitiv. In contrast, the pure approach enforces stricter criteria by requiring complete data from all agencies (Refinitiv, Bloomberg, and S&P500), leading to more balanced but lower correlations. These findings underscore how methodological choices in aggregation influence the variance captured and the dominance of individual data sources in the composite score.

Figure 4.2: Distribution of EAPY_RBS environmental scores for different starting years

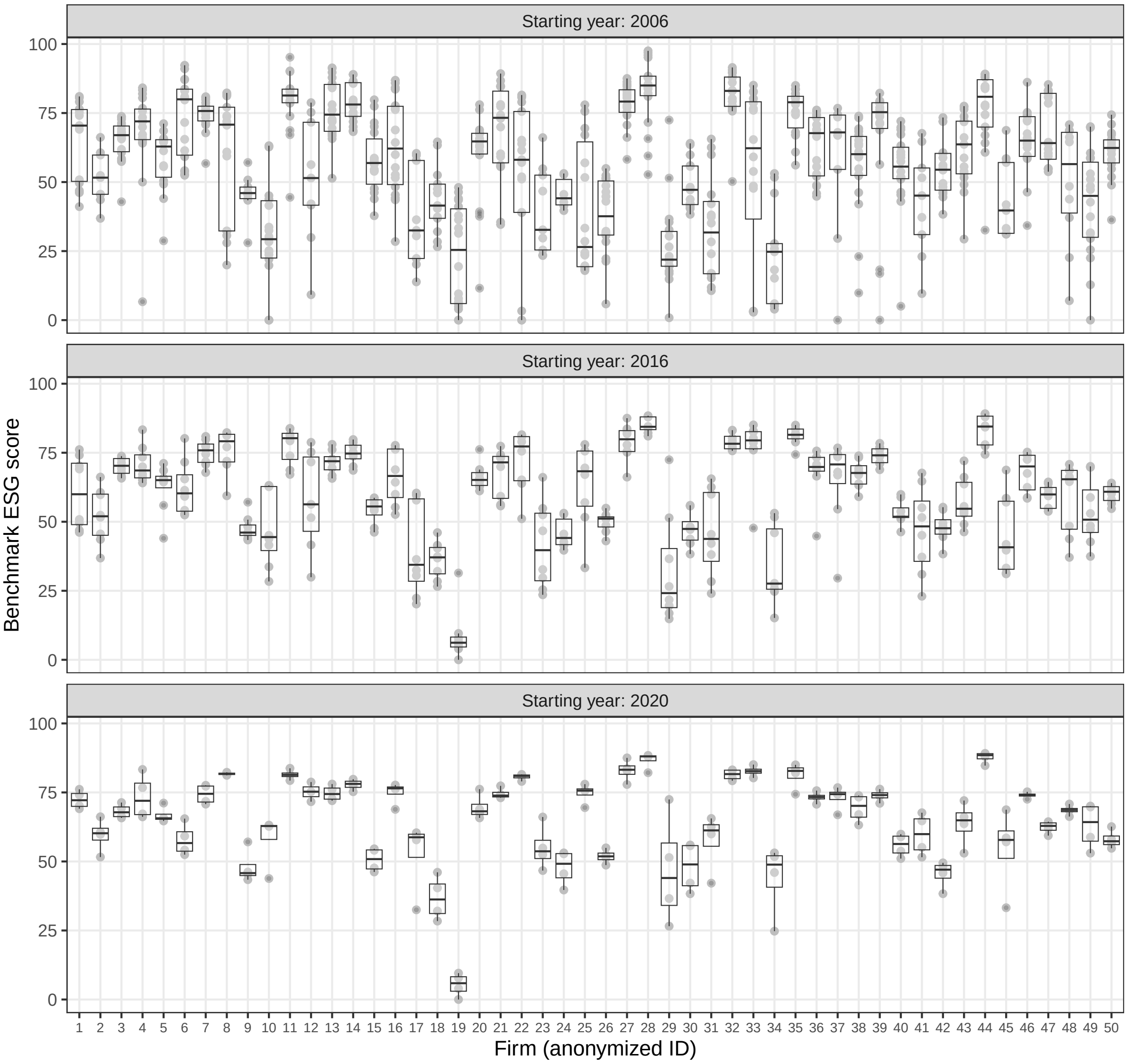


Figure 2 consists of 2 panels of box plots comparing the distribution of EAPY_RBS scores for different starting years: 2006 (top), 2016 (middle) and 2020 (bottom). Each boxplot characterizes the firm-specific distribution of EAPY_RBS score over the period considered by the panel. Each panel represents the EAPY_RBS environmental scores for 50 randomly selected and anonymized firms for clarity. X-Axis: Represents the anonymized firms. Y-Axis: Indicates the EAPY_RBS scores, ranging from 0 to 100.

Figure 4.3: Average standard deviation of composite scores by starting year

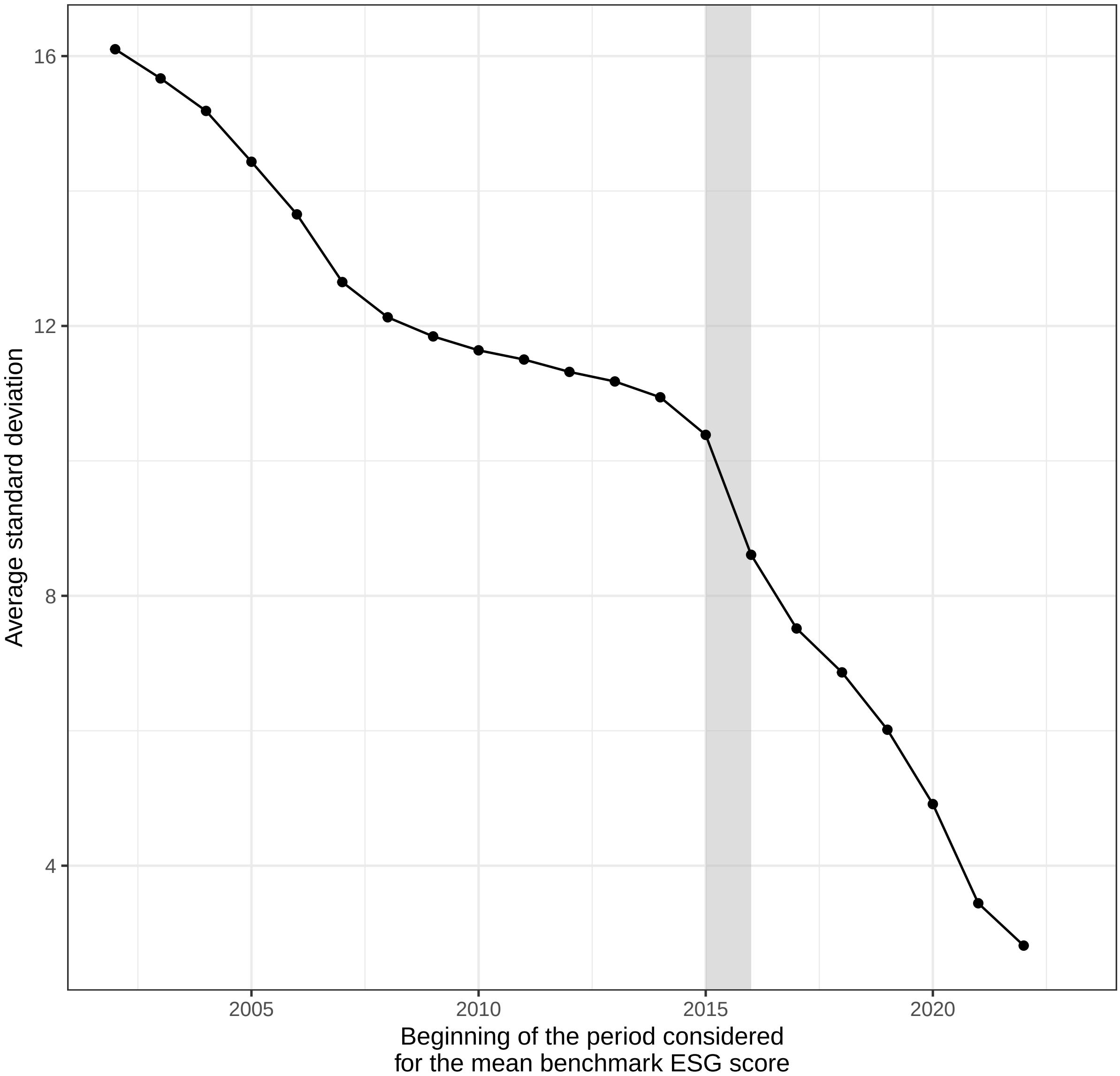


Figure 3 illustrates the average standard deviation of EAPY_RBS, first calculated per firm and then averaged across firms, based on different starting years, ranging from 2002 to 2022. The x-axis denotes the beginning year considered for the mean composite score, while the y-axis represents the average standard deviation of these scores.

Figure 4.4: Distribution Comparison of GPT and RBS Scores

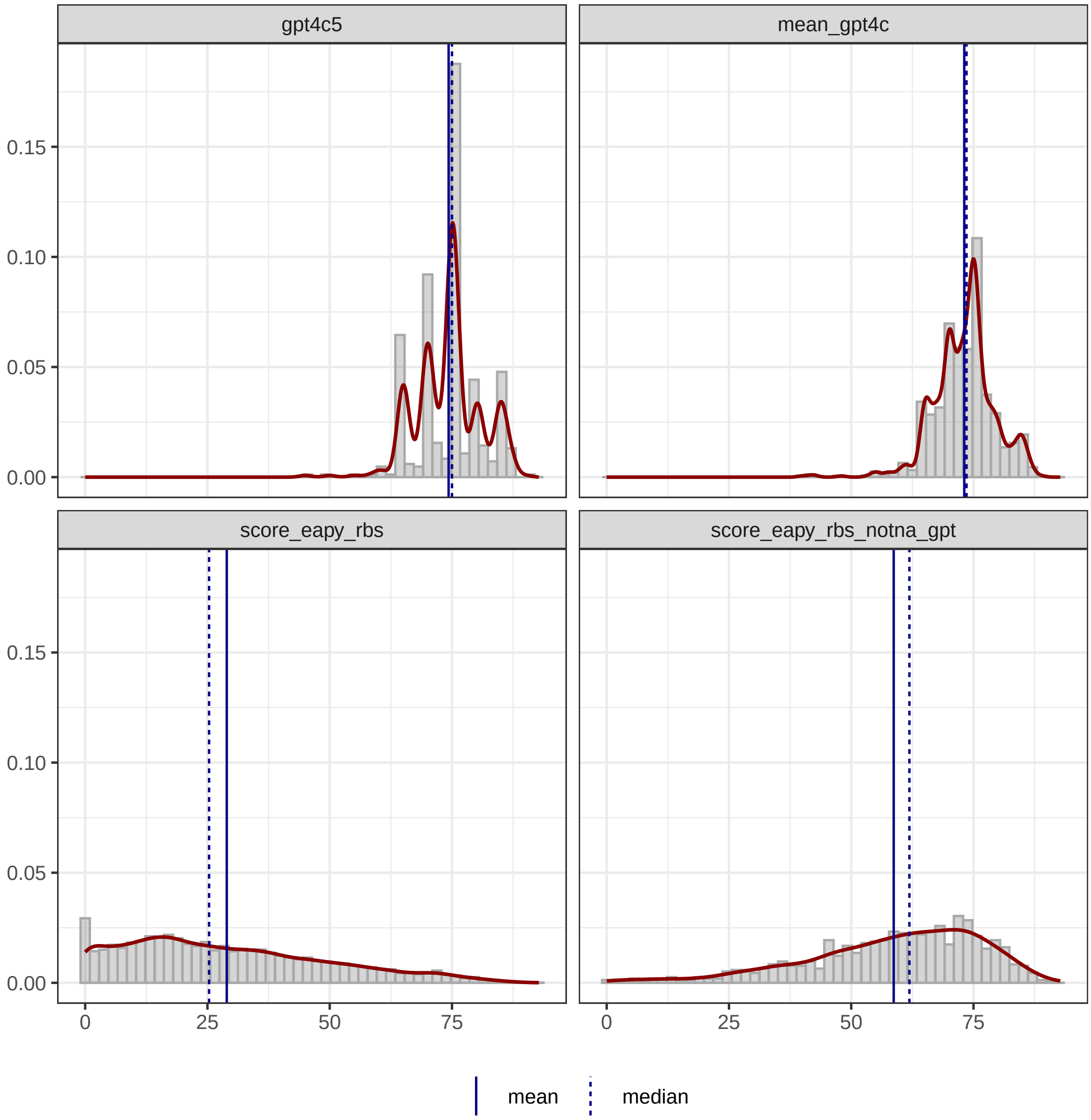


Figure 4 consists of four panels (A, B, C, and D), comparing the distribution of ESG scores generated by GPT and RBS. Each subplot includes a histogram with overlaid density plots, and vertical lines indicating the mean (solid line) and median (dashed line) scores. Figure A (Top Left): Depicts the distribution of environmental scores provided by GPT for a single simulation run. Figure B (Top Right): Illustrates the distribution of the average environmental scores across 10 simulation runs by GPT. Figure C (Bottom Left): Shows the distribution of environmental scores provided by RBS over the period from 2016 to 2023. Figure D (Bottom Right): Displays the distribution of RBS scores, but only for observations where GPT has generated at least one time a score (out of the ten runs).

Figure 4.5: Estimates of measure of association between GPT4 and RBS scores

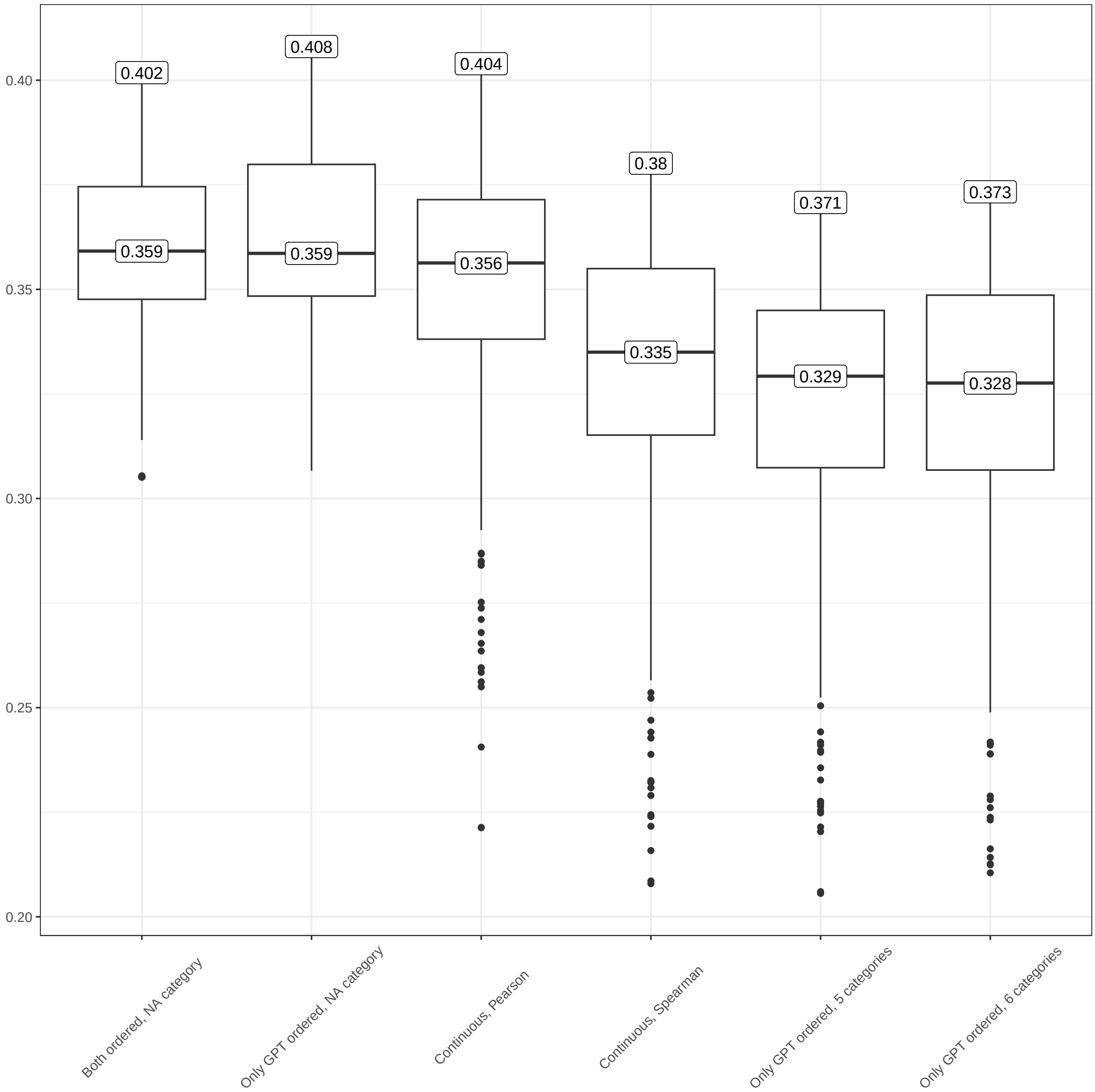


Figure 5 box plots present the distributions of estimates of different measures of association between the 770 possible scores, and different coding of ChatGPT scores. Each boxplot displays the median (thick line), interquartile range (box), and the overall spread of the data (whiskers and outliers). Pearson's r and Spearman's rho are used to quantify the association. When at least one score is encoded as an ordered categorical variable (ordinal scale), the Spearman's rank correlation coefficient is used. The "NA category" indicates whether missing GPT-4 scores are treated as a separate, lowest category. Analysis demonstrates that the inclusion of a missing data category substantially enhance the correlation coefficients (up to 0.408), underscoring the informational value of non-response data, indicating that the inability of GPT-4 to generate a score is informative of poor ESG performance.

Figure 4.6: Comparison of actual GPT categorical scoring with optimal interval scale scoring

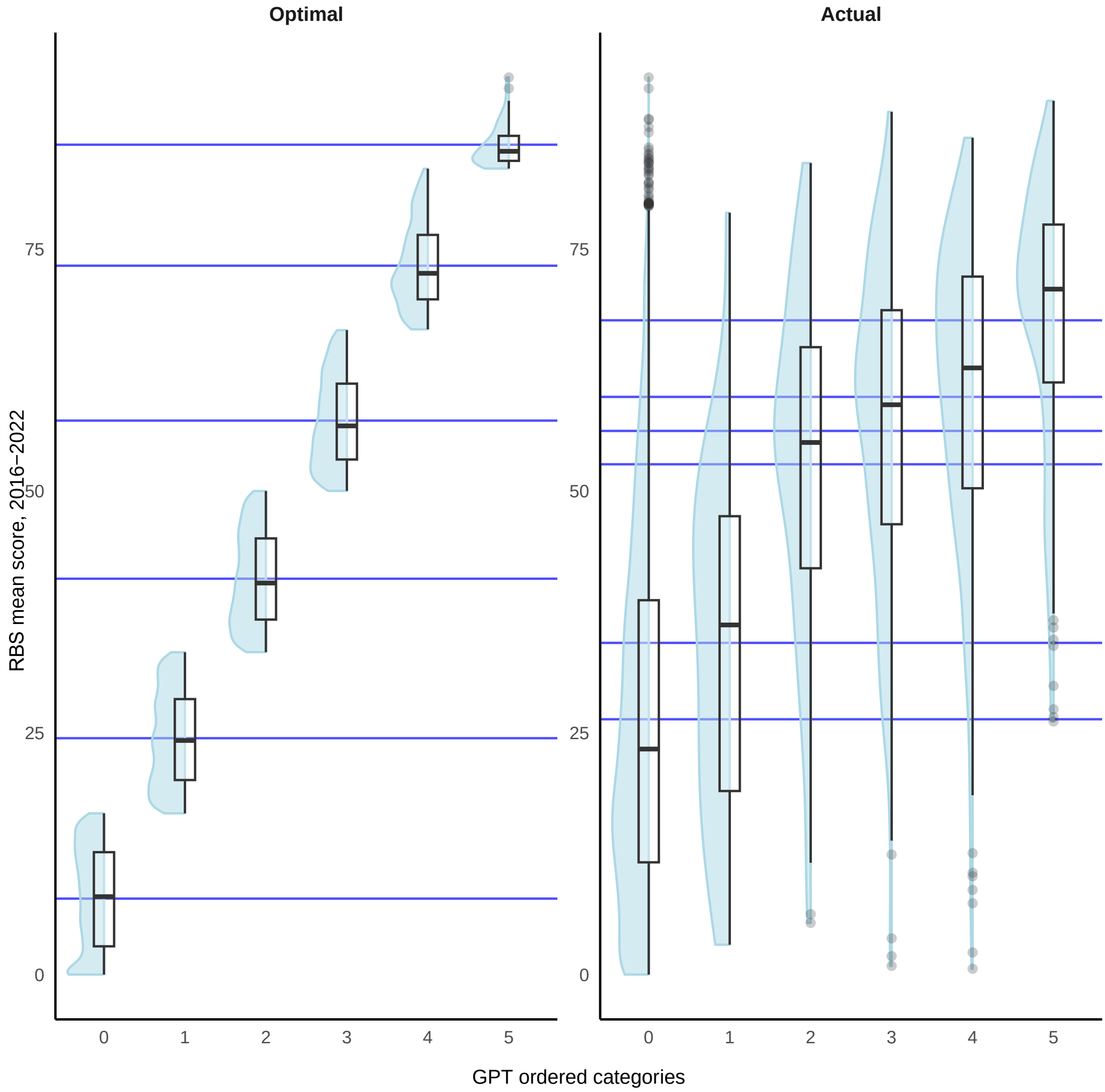
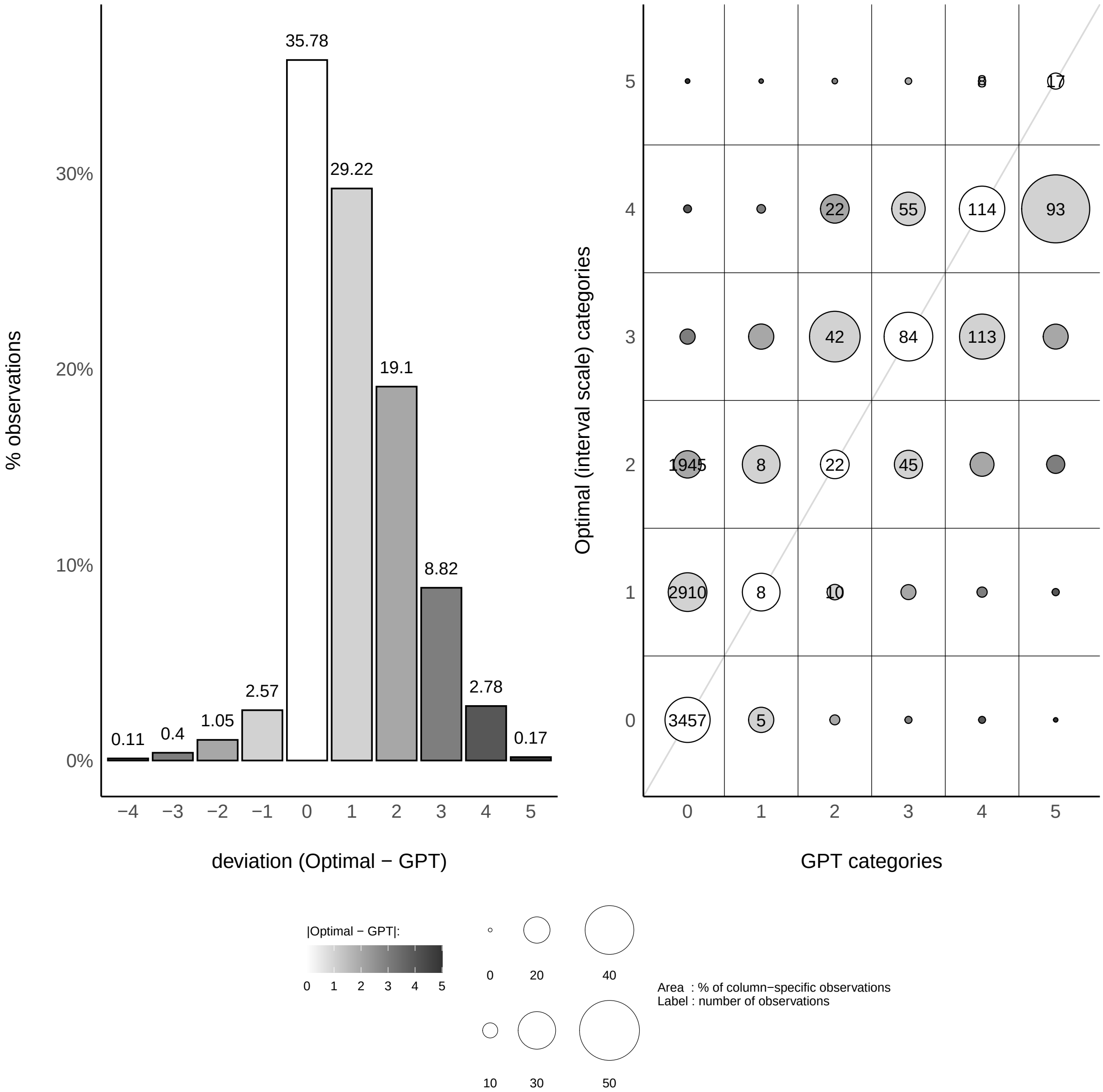


Figure 6 presents combined violin and box plots to compare an optimal interval scale scoring (left panel) with the actual GPT categorical scoring (right panel). The violin plots depict the density distribution of RBS scores within each GPT category. Overlaid black box plots summarize these distributions, showing the median (horizontal line within the box), interquartile range (IQR, the box), and potential outliers (dots outside the whiskers). The left panel (Optimal) shows a smooth progression of scores with clear, equal intervals, reflecting the ideal interval scale. Horizontal lines represent the mean RBS scores by GPT category, highlighting the differences in precision and potential for misclassification in GPT's actual scoring. The right panel (Actual) illustrates the significant overlap and wide dispersion of RBS scores within GPT categories, suggesting high misclassification risk.

Figure 4.7: Misclassification risk assessment: a graphical representation



The left panel represents the share of observations for each level of deviation between Optimal and GPT categories. The bars indicate the share of firms for each level of deviation, with shades representing the severity of the deviation (absolute value). The right panel presents an augmented confusion matrix. Each cell shows the number of firms for a given combination of GPT (columns) and Optimal (rows) categories. The size of the bubbles indicates the percentage of column-specific observations. For example, cell (0, 1) indicates that around 20% of the observations included in the GPT-category 0 (column 0) have been misclassified and should actually belong to optimal category 1 (row 1). This cell contains 2910 observations. Cell (0, 0) indicates that around 40% of the observations in GPT-category 0 are appropriately classified, including 3457 firms.

Figure 4.8: Distribution of differences between GPT and RBS environmental scores

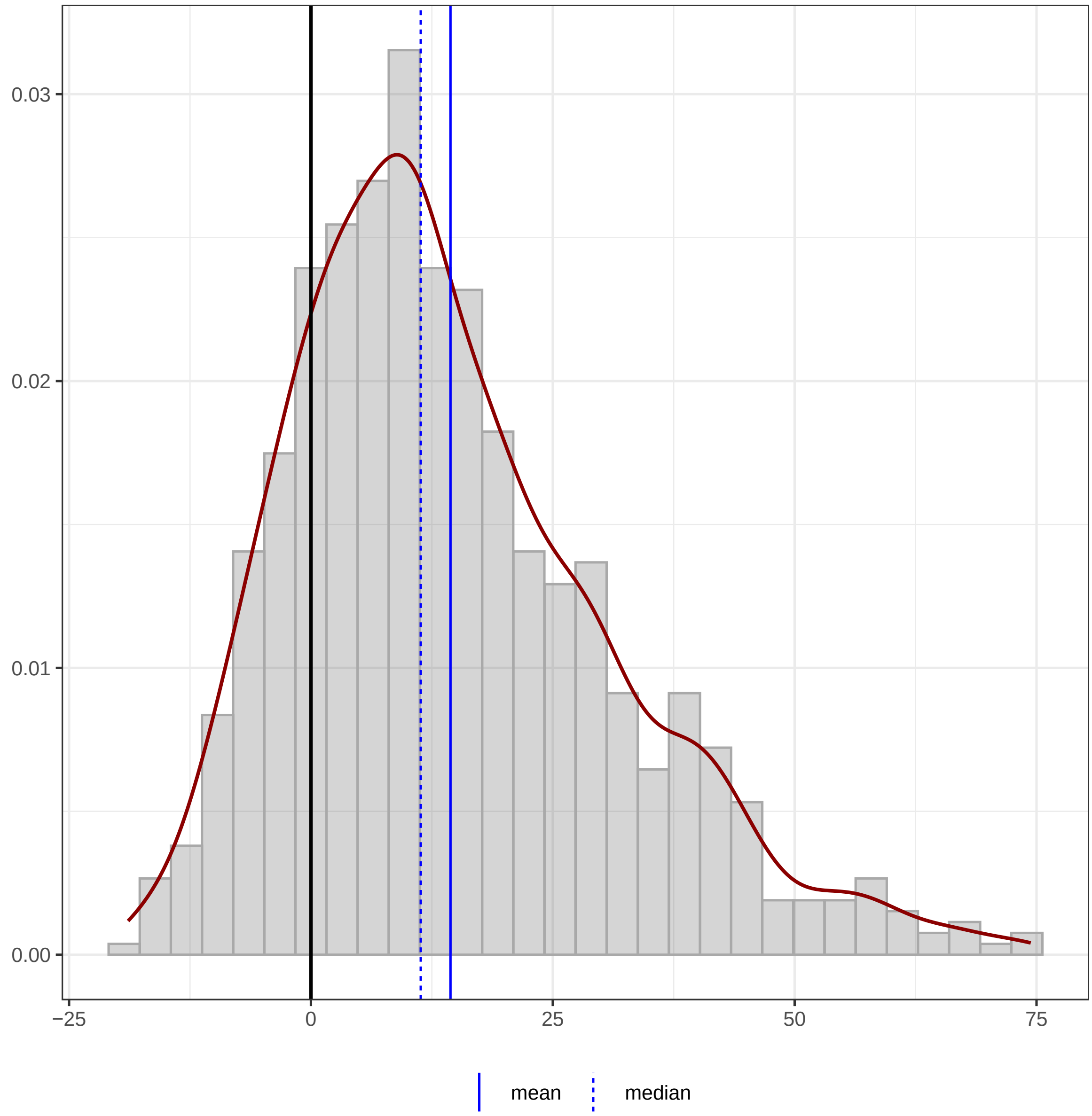


Figure 8 presents a histogram that displays the distribution of the differences between GPT-generated environmental scores and those provided by RBS. The x-axis represents the difference in scores (GPT minus RBS), while the y-axis shows the density of occurrences. The solid second line indicates the mean difference, and the dashed line represents the median difference.

Figure 4.9: Determinants of ESG performance

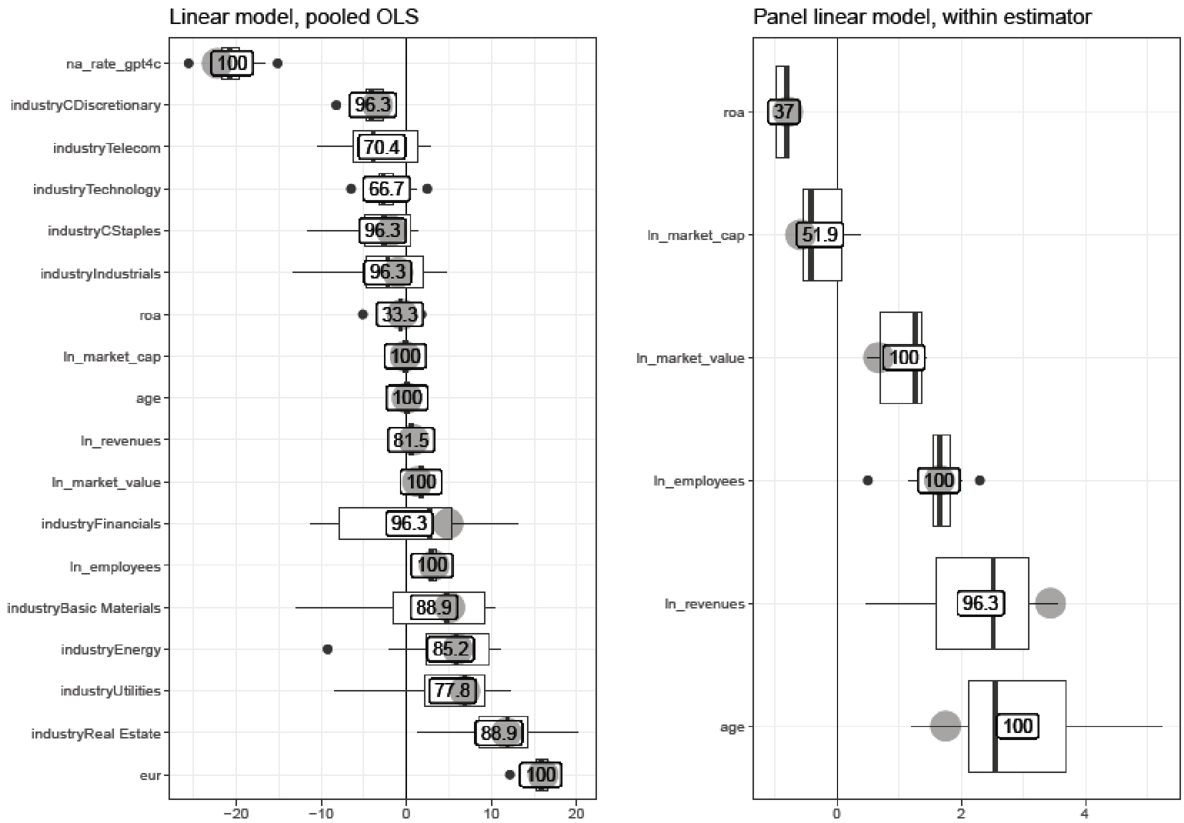


Figure 9 presents the results of the determinants of ESG performance. The left panel uses a linear regression model, estimated by (pooled) OLS. The right panel uses a panel linear regression model, estimated via within estimators. The box plots display the distributions of the coefficient estimates across the 27 different model specifications [(6 possible aggregation methods * 4 possible score combinations) +3 single agency scores]. The numbers within the boxes indicate the percentage of models for which the corresponding coefficient is statistically significant (p-value < 0.05). The grey circles highlight the results for the selected baseline model and period, EAPY 2016.

Figure 4.10: Analysis of factors influencing the environmental scores generation and precision

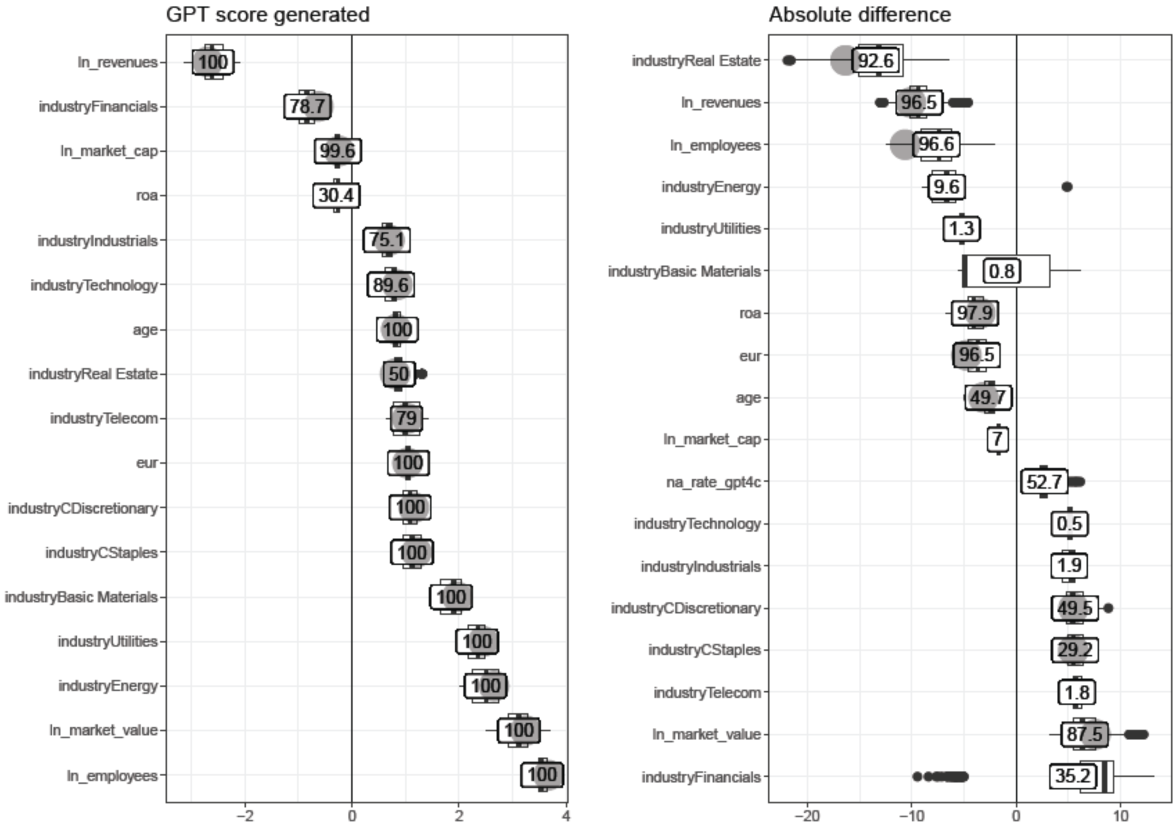


Figure 10 presents the results from two different models analyzing various factors influencing the environmental scores generated by GPT-4 in comparison to benchmark scores. The box plots display the distributions of the coefficient estimates across the 770 benchmark scores. The numbers within the boxes indicate the percentage of models for which the corresponding coefficient is statistically significant (p-value < 0.05). The grey circles highlight the results for the selected baseline model and period, EAPY RBS 2016. Model (1) – GPT score generated (Binary): identifies the factors that affect whether GPT-4 generates a score for a given company, with 0 indicating a missing score and 1 indicating a generated score. Model (2) – Absolute difference: identifies the factors that affect the absolute value of the difference between GPT-4 and benchmark scores.

Figure 4.11: Analysis of factors influencing the environmental scores overestimation probability and intensity

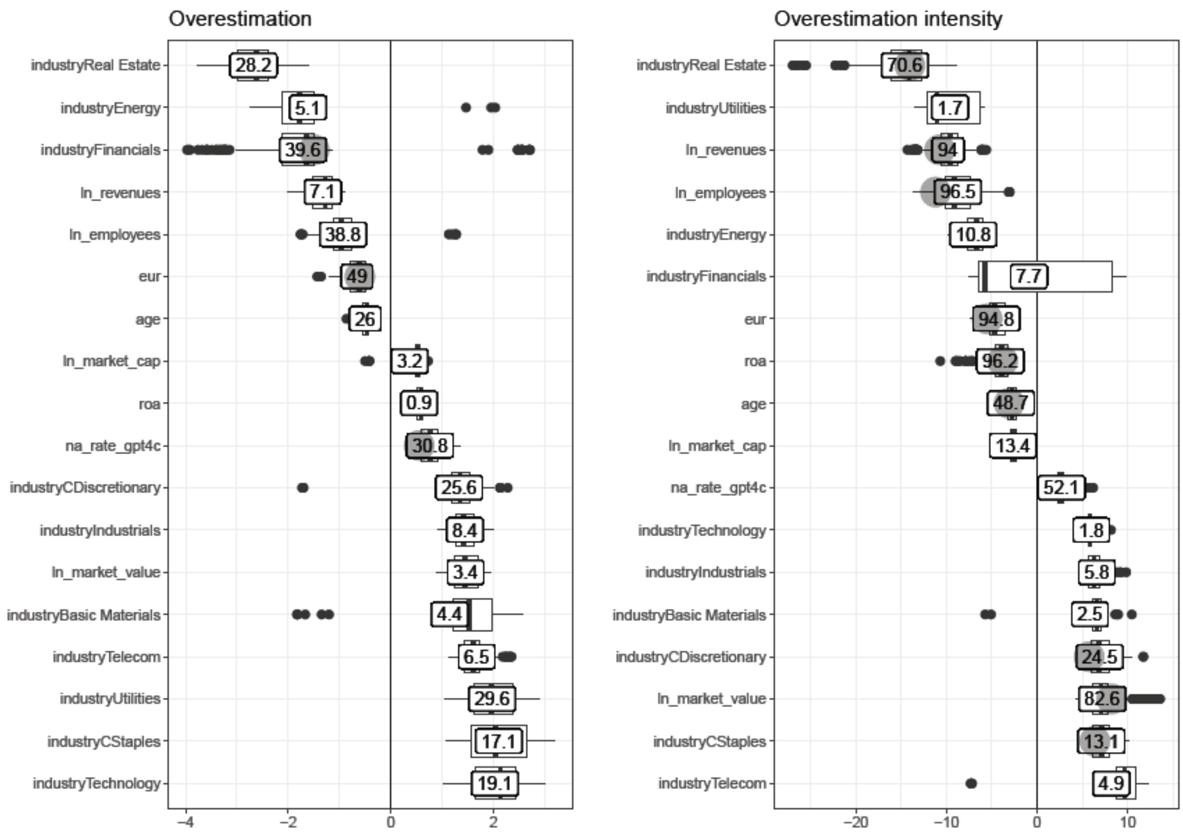


Figure 11 presents the results from two different models analyzing various factors influencing the environmental scores generated by GPT-4 in comparison to the benchmark scores. The box plots display the distributions of the coefficient estimates across the 770 different benchmark scores. The numbers within the boxes indicate the percentage of models for which the corresponding coefficient is statistically significant (p-value < 0.05). The grey circles highlight the results for the selected baseline model and period, EAPY RBS 2016. Model (1) – Overestimation (Binary): indicates whether the GPT-4 score is higher than the benchmark score, with 0 if (GPT − benchmark score) ≤ 0 and 1 if GPT $>$ benchmark score. Model (2) – Overestimation Intensity (Level): quantifies the level of overestimation by GPT-4, measuring the difference |GPT−benchmark score| where Overestimation (Binary) = 1.

Table 4.2: Summary table of different score aggregation methods

Model	Combination type	Methodology	Approach Description
EA	**Opportunistic:** Allow for the computation of composites even when not all agencies report in a given year, maximizing data utilization but potentially introducing variability in the weighting of agency contributions.	Simple Average	Yearly average of available agency scores.
EAP		PCA-Weighted Average (Pooled)	Pools data across years to derive weights from the first principal component loadings computed on the entire sample, ensuring consistency. When scores are missing for a firm-year combination, a simple average of available scores is used to avoid missing values, maintaining coverage across all observations.
EAPY		PCA-Weighted Average (Yearly)	PCA-derived weights applied to agency scores for each year, with PCA recalculated annually. If there are not enough non-missing values to estimate PCA factor loadings reliably, a simple mean of the available scores is calculated to ensure a composite score is still produced. If only one agency has scores, its values are used directly, and if no agency has scores, the composite score is set to NA.
PEA	**Pure:** Only compute averages when all selected agencies provide data.	Simple Average	Yearly average of agency scores, calculated only when all selected agencies provide data.
PEAP		PCA-Weighted Average (Pooled)	Pure Linear combination of agency scores, with weights derived from PCA loadings calculated on the entire pooled sample. Missing firm-year scores result in the composite score being NA, as the linear combination includes NA values, ensuring strict adherence to the PCA weighting scheme.
PEAPY		PCA-Weighted Average (Yearly)	PCA-derived weights applied annually, calculated only when all selected agencies provide data. This makes the measurement concept "pure," as it relies entirely on complete data. However, this strict criterion leads to a significant increase in missing values.
A	N/A	Simple Average	First averages agency scores over the selected period, then combines them into a composite using simple averages.
P	N/A	PCA-Weighted Average	First averages agency scores over the selected period, then combines them into a composite using PCA-derived weights.

Table 2 summarizes all the aggregation methods applied to all the possible combination of agencies (RBS, RB, RS, BS). These approaches reflect different trade-offs between data completeness and methodological consistency and conceptual purity. PCA loadings are normalized so that the sum of weights equals 1. For example, EAP_RBS model in 2004 would take the value of Refinitiv, while the PEAP_RBS model in 2004 would not yield a score since not all agencies reported scores; this ensures purity but reduces the number of usable observations, especially noticeable in years prior to 2015 when Bloomberg and S&P500 scores are largely absent.

Table 4.3: Comparison of availability rates for GPT-4 scores under restricted and unrestricted prompts

Source	Scope	Availability (%)
GPT 4 (unrestricted)	Unique run	15.69
GPT 4 (restricted)	Mean (runs)	7.89
GPT 4 (restricted)	Run 10	6.38
GPT 4 (restricted)	Run 7	4.32
GPT 4 (restricted)	Run 5	4.25
GPT 4 (restricted)	Run 9	4.23
GPT 4 (restricted)	Run 1	4.18
GPT 4 (restricted)	Run 8	4.18
GPT 4 (restricted)	Run 4	4.17
GPT 4 (restricted)	Run 3	4.13
GPT 4 (restricted)	Run 6	4.12
GPT 4 (restricted)	Run 2	4.10

Table 3 compares the availability percentage of ESG scores generated by GPT-4 using restricted and unrestricted prompts. The unrestricted prompt refers to a unique run without any specific limitations, while the restricted prompt involves multiple runs where GPT-4 is explicitly instructed to state "No specific information available" when it lacks information for any category.

Table 4.4: Correlation Between environmental scores from traditional agencies

couple	Pearson(panel)	Spearman(panel)	n(panel)	Pearson(med.)	Spearman(med.)	n(med.)
Refinitiv, Bloomberg	0.503	0.558	43513	0.509	0.577	8464
Refinitiv, SP500	0.588	0.587	32755	0.624	0.593	7159
Bloomberg, SP500	0.449	0.447	31107	0.476	0.464	6784

Table 4 presents the correlation coefficients (Pearson and Spearman) among ESG scores from Refinitiv, Bloomberg, and S&P500. The table includes correlations of the longitudinal version of the dataset (i.e. firm-year), and the median value of the correlation coefficient estimates calculated across datasets with varying starting years (med.), as well as the number of observations (n). Panel-based Pearson correlations range from 0.449 (Bloomberg and S&P500) to 0.588 (Refinitiv and S&P500), while median-based Pearson correlations range from 0.476 to 0.624. Spearman correlations follow a similar pattern. These values reflect moderate alignment, with stronger relationships observed between Refinitiv and S&P500 compared to Bloomberg and S&P500.

8 Appendices

Appendix A - Prompt engeneering

Following the construction of the environmental pillar of Refinitiv ESG scores:

Table 4.5: Refinitiv/Asset4 ESG rating elements; source: Refinitiv (2022)

Pillars	Catagories	Themes	Data points	Weight method
Environmental	Emmission	Emissions	TR.AnalyticCO2	Quant industry median
		Waste	TR.AnalyticTotalWaste	Quant industry median
		Biodiversity*		
		Environmental management systems*		
	Innovation	Product innovation	TR.EnvProducts	Transparency weights
		Green revenues, research and development (R&D) and capital expenditures (CapEx)	TR.AnalyticEnvRD	Quant industry median
	Resource use	Water	TR.AnalyticWaterUse	Quant industry median
		Energy	TR.AnalyticEnergyUse	Quant industry median
		Sustainable packaging*		
		Environmental supply chain*		
Social	Community	Equally important to all industry groups, hence a median weight of five is assigned to all		Equally important to all industry groups
	Human rights	Human rights	TR.PolicyHumanRights	Transparency weights
	Product responsibility	Responsible marketing	TR.PolicyResponsibleMarketing	Transparency weights
		Product quality	TR.ProductQualityMonitoring	Transparency weights
		Data privacy	TR.PolicyDataPrivacy	Transparency weights
	Workforce	Diversity and inclusion	TR.WomenEmployees	Quant industry median
		Career development and training	TR.AvgTrainingHours	Transparency weights
		Working conditions	TR.TradeUnionRep	Quant industry median
		Health and safety	TR.AnalyticLostDays	Transparency weights
Governance	CSR strategy	CSR strategy	Data points in governance category and governance pillar	Count of data points in each governance category/all data points in governance pillar
		ESG reporting and transparency		
	Management	Structure (independence, diversity, committees)	Data points in governance category and governance pillar	Count of data points in each governance category/all data points in governance pillar
		Compensation		
	Shareholders	Shareholder rights	Data points in governance category and governance pillar	Count of data points in each governance category/all data points in governance pillar
		Takeover defenses		

*No data points available that may be used as a proxy for ESG magnitude/materiality

https://www.lseg.com/content/dam/data-analytics/en_us/documents/methodology/lseg-esg-scores-methodology.pdf

The prompt is constructed as follows:

Figure 4.12: Prompt construction

```
f"Please provide a detailed evaluation of {row['company']}'s sustainability
"and innovation efforts based on verified and available information as of the"
"latest update, specifically focusing on:\n\n"
"- Emission: Documented evidence of the company's actions and effectiveness"
"in reducing emissions, waste management, biodiversity preservation, and the"
"implementation of environmental management systems.\n"
"- Innovation: Verified activities and investments related to product "
" innovation, green revenues, research and development (R&D) investments,"
" and capital expenditures (CapEx) in the field of sustainability.\n"
"- Resource Use: Concrete examples of the company's practices regarding water
"and energy usage, sustainable packaging, and promoting environmental "
"responsibility in its supply chain.\n\n"
"For each category with available information, provide a score "
"on a scale of 0 to 100, reflecting the company's achievements,
" ongoing challenges, and potential areas for improvement."
"Summarize the findings with an overall total score out of 100, calculated as
"the mean of the available category scores.\n\n"
"Structure your response as follows:\n\n"
"- Emission: [If no information is available, omit this section.]\n"
"- Innovation: [If no information is available, omit this section.]\n"
"- Resource Use: [If no information is available, omit this section.]\n"
"- Total Score: /100 [Calculated only from available scores.]\n\n"
"Note: If you have no information available for any of the categories,"
" explicitly state 'No specific information available' for each."
"Do not provide a general industry assessment or speculate on practices based on"
"industry norms. This response requires information specific to the named company"
"and should not include assumptions or generalizations about industry practices."
```

Figure 12 presents the prompt that was submitted to OpenAI's Batch API to generate a detailed evaluation of a company's sustainability efforts based on specific environmental factors such as emissions reduction, waste management, and innovation in sustainable products. The prompt follows the categories of Refinitiv's environmental pillar, ensuring alignment with established ESG frameworks. The prompt outlines a structured approach for ChatGPT to assess these areas and score them, highlighting the model's use in complementing traditional ESG assessments with qualitative insights, while ensuring accuracy through conditions that prevent data hallucinations.

Appendix B - Variables used in the study

Table 4.6: Variables used in the study

Variable	Measure	Description
Firm Age	T - Date of incorporation	The number of years a firm has been in operation, found by calculating the years since the date of incorporation (Datastream: WC18273)
Financial Performance	ROA (Return on Assets)	Measures the company's profitability by dividing total assets by net incomes. (Datastream: WC02999 / WC01751)
Firm Size	Number of Employees	Proxy for firm size representing the number of both full and part time employees of the company. (Datastream: WC07011)
	Revenues	Proxy for firm size representing gross sales and other operating revenue less discounts, returns and allowances. (Datastream: WC01001)
	Market Value	Proxy for firm size the consolidated Market Value of a company expressed in millions, calculated by summing the market values of all eligible securities issued by the company. (Datastream: MVC)
	Market Capitalization	Proxy for firm size representing the total market value of the firm's outstanding shares, measured as Market Price-Year End * Common Shares Outstanding. (Datastream: WC08001)
Geographic Location	Europe, Country	The country of the legal domicile of a company's head office. (Datastream: CODON)
Sector/Industry	NACE R2	FTSE Russell Industry Classification Benchmark (ICB) Industry level classification name (Datastream: ICBIN)

Table 6 outlines the variables analyzed in the study, including their respective measures and descriptions. The variables encompass firm age, financial performance, firm size, geographic location, and sector/industry classification, with specific data sources cited for each measure.

Appendix C - Robustness of inference

Figure 4.13: Analysis of factors influencing the environmental scores generation and precision with bootstrapped standard errors

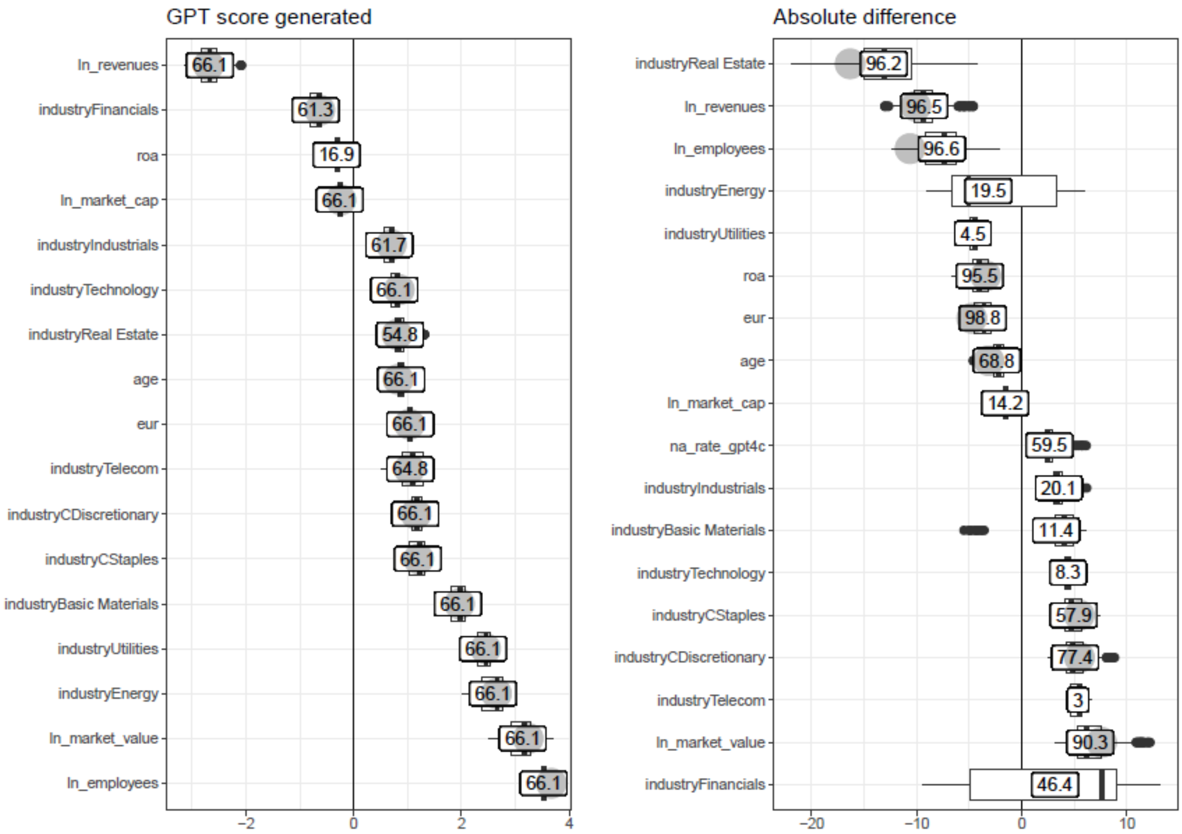


Figure 13 presents a bootstrapped version of the coefficient estimates of Figure 10. This approach addresses the additional variability introduced by PCA-derived weights, ensuring more reliable inference. Bootstrapping is generally more robust against the violation of standard assumptions used in asymptotic (traditional) statistical tests. It can provide a more accurate representation of the uncertainty in parameter estimates when the underlying data conditions are complex or not ideal for traditional methods. The updated figure now includes these bootstrapped results, highlighting a lower percentage of significant estimates compared to the standard inference.

Figure 4.14: Analysis of factors influencing the environmental scores overestimation probability and intensity with bootstrapped standard errors

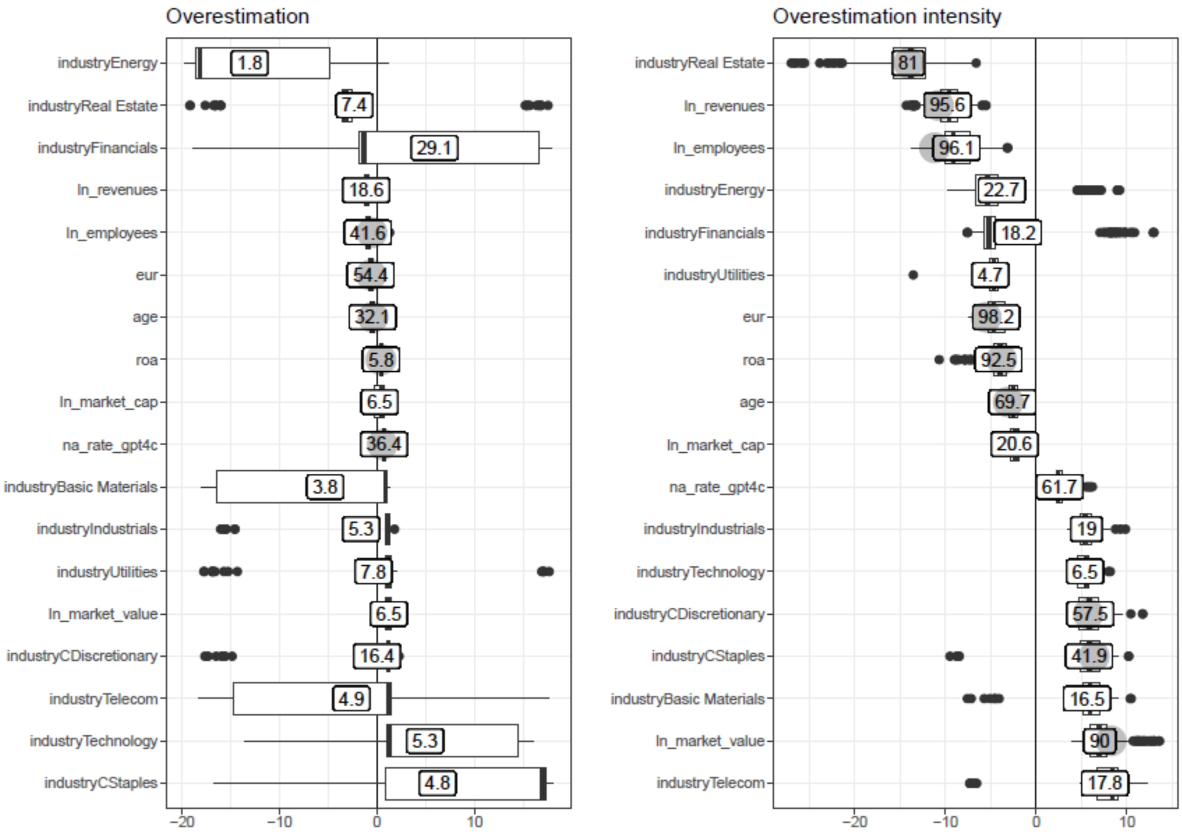


Figure 14 presents a bootstrapped version of the coefficient estimates of Figure 11. This approach addresses the additional variability introduced by PCA-derived weights, ensuring more reliable inference. Bootstrapping is generally more robust against the violation of standard assumptions used in asymptotic (traditional) statistical tests. It can provide a more accurate representation of the uncertainty in parameter estimates when the underlying data conditions are complex or not ideal for traditional methods. The updated figure now includes these bootstrapped results, highlighting a lower percentage of significant estimates compared to the standard inference.

9 References

Aguinis, H., & Glavas, A. (2012). What we know and don't know about corporate social responsibility: A review and research agenda [publisher: Sage Publications Sage CA: Los Angeles, CA]. *Journal of management*, *38*(4), 932–968. https://doi.org/10.1177/0149206311436079 (cit. on p. 175).

Ayinde, L., Wibowo, M. P., Ravuri, B., & Emdad, F. B. (2023). Chatgpt as an important tool in organizational management: A review of the literature [publisher: SAGE Publications Sage UK: London, England]. *Business Information Review*, *40*(3), 137–149. https://doi.org/10.1177/02663821231187991 (cit. on pp. 172, 173).

Bandalos, D. L. (2018, January). *Measurement Theory and Applications for the Social Sciences* [Google-Books-ID: caxCDwAAQBAJ]. Guilford Publications. (Cit. on p. 185).

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big?, 610–623 (cit. on p. 188).

Berg, F., Koelbel, J. F., & Rigobon, R. (2022). Aggregate confusion: The divergence of ESG ratings [publisher: Oxford University Press]. *Review of Finance*, *26*(6), 1315–1344. https://doi.org/10.1093/rof/rfac033 (cit. on p. 171).

Billio, M., Costola, M., Hristova, I., Latino, C., & Pelizzon, L. (2021). Inside the ESG ratings:(Dis)agreement and performance [publisher: Wiley Online Library]. *Corporate Social Responsibility and Environmental Management*, *28*(5), 1426–1445. https://doi.org/10.1002/csr.2177 (cit. on p. 171).

Binns, R., Van Kleek, M., Veale, M., Lyngs, U., Zhao, J., & Shadbolt, N. (2018). 'It's Reducing a Human Being to a Percentage' Perceptions of Justice in Algorithmic Decisions, 1–14 (cit. on p. 175).

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/https://doi.org/10.1002/smj.2407 (cit. on p. 172).

Christensen, D. M., Serafeim, G., & Sikochi, 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 (cit. on pp. 172, 175).

Clarkson, P. M., Li, Y., Richardson, G. D., & Vasvari, F. P. (2008). Revisiting the relation between environmental performance and environmental disclosure: An empirical analysis [publisher: Elsevier]. *Accounting, organizations and society*, *33*(4-5), 303–327. https://doi.org/10.1016/j.aos.2007.05.003 (cit. on p. 175).

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 (cit. on p. 175).

Diebecker, J., Rose, C., & Sommer, F. (2017). Corporate sustainability performance over the firm life cycle: Levels, determinants, and the impact on accounting performance. *Determinants, and the Impact on Accounting Performance (December 8, 2017)* (cit. on pp. 175, 176).

Dilling, P. F. (2010). Sustainability reporting in a global context: What are the characteristics of corporations that provide high quality sustainability reports an empirical analysis. *International Business & Economics Research Journal (IBER)*, *9*(1). https://doi.org/10.19030/iber.v9i1.505 (cit. on p. 176).

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 (cit. on p. 172).

Drobetz, W., Merikas, A., Merika, A., & Tsionas, M. G. (2014). Corporate social responsibility disclosure: The case of international shipping [publisher: Elsevier]. *Transportation Research Part E: Logistics and Transportation Review*, *71*, 18–44. https://doi.org/10.1016/j.tre.2014.08.006 (cit. on p. 176).

Eccles, R. G., Ioannou, I., & Serafeim, G. (2014). The impact of corporate sustainability on organizational processes and performance [publisher: Informs]. *Management* *science*, *60*(11), 2835–2857. https://doi.org/10.1287/mnsc.2014.1984 (cit. on p. 175).

Garcia, A. S., Mendes-Da-Silva, W., & Orsato, R. J. (2019). Corporate Sustainability, Capital Markets, and ESG Performance [DOI: 10.1007/978-3-319-91911-9_13]. In *Individual Behaviors and Technologies for Financial Innovations* (pp. 287–309). https://link.springer.com/10.1007/978-3-319-91911-9_13 (cit. on p. 176).

Gjølberg, M. (2009). The origin of corporate social responsibility: Global forces or national legacies? [publisher: Oxford University Press]. *Socio-economic review*, *7*(4), 605–637. https://doi.org/10.1093/ser/mwp017 (cit. on p. 176).

Helbig, R., von Höveling, S., Solsbach, A., & Marx Gómez, J. (2021). Strategic analysis of providing corporate sustainability open data. *Intelligent Systems in Accounting, Finance and Management*, *28*(3), 195–214. https://doi.org/10.1002/isaf.1501 (cit. on p. 172).

Hoffmann, J., Borgeaud, S., Mensch, A., Buchatskaya, E., Cai, T., Rutherford, E., & Sifre, L. (2022). Training Compute-Optimal Large Language Models [arXiv:2203.15556 [cs]]. https://doi.org/10.48550/arXiv.2203.15556 (cit. on p. 174).

Hoogendoorn, B., Van Der Zwan, P., & Thurik, R. (2019). Sustainable Entrepreneurship: The Role of Perceived Barriers and Risk. *Journal of Business Ethics*, *157*(4), 1133–1154. https://doi.org/10.1007/s10551-017-3646-8 (cit. on p. 172).

Jennifer Ho, L., & Taylor, M. E. (2007). An empirical analysis of triple bottomline reporting and its determinants: Evidence from the United States and Japan [publisher: Wiley Online Library]. *Journal of International Financial Management & Accounting*, *18*(2), 123–150. https://doi.org/10.1111/j.1467-646X.2007.01010.x (cit. on pp. 175, 176).

Kim, A., Muhn, M., & Nikolaev, V. V. (2024). Bloated disclosures: Can ChatGPT help investors process information? *Chicago Booth Research Paper*, (23-07), 2023–59 (cit. on pp. 172, 173).

Kokina, J., & Davenport, T. H. (2017). The emergence of artificial intelligence: How automation is changing auditing. *Journal of Emerging Technologies in Accounting*, *14*(1), 115–122. https://doi.org/10.2308/jeta-51730 (cit. on p. 172).

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 (cit. on p. 172).

Krueger, P., Sautner, Z., Tang, D. Y., & Zhong, R. (2021). The Effects of Mandatory ESG Disclosure Around the World [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1475-679X.12548]. *Journal of Accounting Research*, *n/a*. https://onlinelibrary-wiley-com.proxy.bnl.lu/doi/abs/10.1111/1475-679X.12548. https://doi.org/10.1111/1475-679X.12548 (cit. on p. 175).

Marquis, C., & Qian, C. (2014). Corporate social responsibility reporting in China: Symbol or substance? [publisher: Informs]. *Organization science*, *25*(1), 127–148. https://doi.org/10.1287/orsc.2013.0837 (cit. on p. 176).

Matten, D., & Moon, J. (2008). “Implicit” and “Explicit” CSR: A Conceptual Framework for a Comparative Understanding of Corporate Social Responsibility [publisher: Academy of Management]. *Academy of Management Review*, *33*(2), 404–424. https://doi.org/10.5465/amr.2008.31193458 (cit. on p. 176).

Maynez, J., Narayan, S., Bohnet, B., & McDonald, R. (2020). On faithfulness and factuality in abstractive summarization. *arXiv preprint arXiv:2005.00661* (cit. on pp. 182, 188).

MerklDavies, D. M., Brennan, N. M., & McLeay, S. J. (2011). Impression management and retrospective sensemaking in corporate narratives: A social psychology perspective [publisher: Emerald Group Publishing Limited]. *Accounting, Auditing & Accountability Journal*, *24*(3), 315–344. https://doi.org/10.1108/09513571111124036 (cit. on p. 175).

Michelon, G., Pilonato, S., & Ricceri, F. (2015). Csr reporting practices and the quality of disclosure: An empirical analysis [publisher: Elsevier]. *Critical perspectives on accounting*, *33*, 59–78. https://doi.org/10.1016/j.cpa.2014.10.003 (cit. on p. 171).

Morhardt, J. E. (2010). Corporate social responsibility and sustainability reporting on the Internet. *Business Strategy and the Environment*, *19*(7), 436–452. https://doi.org/10.1002/bse.657 (cit. on p. 175).

Ni, J., Bingler, J., Colesanti-Senni, C., Kraus, M., Gostlow, G., Schimanski, T., & Webersinke, N. (2023). Paradigm Shift in Sustainability Disclosure Analysis: Empowering Stakeholders with CHATREPORT, a Language Model-Based Tool. *arXiv preprint arXiv:2306.15518* (cit. on pp. 172, 173).