

The Language of Greenwashing: SDG Omission and Opportunity-oriented Environmental Tone as Alert Metrics in Green Bond Disclosures

Andrea Nicolodi^{1,2}, Sandra Paterlini¹, Monica Gentile³, Vincenzo
Foglia Manzillo³, and Gianluca Vittorioso³

¹University of Trento, Italy

²Free University of Bolzano-Bozen, Italy

³CONSOB, Rome, Italy

December 18, 2025

DO NOT CIRCULATE WITHOUT AUTHORS' PERMISSION*

*The views expressed in this article are those of the authors and do not necessarily reflect those of CONSOB.

Abstract

Green bonds play a central role in sustainable finance, yet concerns about greenwashing raise questions about the credibility of issuers’ sustainability disclosures. Using dictionary-based methods and domain-specific BERT transformer models, this paper proposes two greenwashing alert metrics and investigates their performance by analyzing sustainability reports of European corporate green bond issuers from 2019–2023.

First, we develop the SDGs Omission Index (SDGOI), which measures the discrepancy between Sustainable Development Goals (SDGs) declared in green bond frameworks and those subsequently reported in sustainability disclosures. We show that the use of SDG-specific language is associated with broader SDG coverage, whereas generic environmental language is not.

Second, we introduce the Environmental Sentiment Metric (ESM), capturing opportunity-oriented environmental sentiment. We find that a more opportunity-oriented tone is positively related to SDG omission, ESG controversies, and greenwashing allegations.

Together, SDGOI and ESM provide interpretable, disclosure-based indicators that enhance the detection of greenwashing risk in the green bond market.

JEL classifications: G34, G38, J33, K22, M52.

Keywords: greenwashing alert, green bonds, AI, NLP, SDG omission, sentiment analysis.

1 Introduction

Green bonds are pivotal in financing the transition to a sustainable economy by channeling capital into environmentally beneficial projects such as renewable energy, clean transportation, and climate resilience, in line with the 2015 Paris Agreement and the United Nations Sustainable Development Goals (SDGs). To ensure transparency and impact, green bond proceeds should be allocated to well-defined eligible projects. In 2024, green bonds emerged as the dominant category within the sustainable finance market, with issuances reaching USD 381 billion in the corporate sector and USD 257 billion in the public sector, according to the OECD’s 2025 Global Debt

Report.¹

Until the adoption of Regulation (EU) 2023/2631 on European green bonds, the EU lacked a unified standard for certifying green financial instruments. Issuers relied on voluntary frameworks, such as the Green Bond Principles (ICMA) and the Climate Bonds Initiative (CBI), which provided guidance but no enforceable compliance. While these frameworks are widely recognized, rising concerns about greenwashing highlight that some issuers may use the green label without accurately reflecting sustainability performance or delivering real environmental impact. Common signs of greenwashing include vague sustainability claims, selective disclosure, and weak alignment with recognized frameworks. In the absence of robust and harmonized verification mechanisms, companies may make misleading claims about their alignment with the SDGs, often through non-binding references that are not fully integrated into the firm’s Environmental, Social and Governance (ESG) strategy. One frequent manifestation is the selective disclosure of SDG alignment while omitting goals that might reveal trade-offs, negative externalities, or insufficient ambition. From a corporate social responsibility perspective, such practices undermine stakeholder trust, distort capital allocation, and weaken the credibility of sustainability reporting as a governance tool.

For many stakeholders, the primary observable evidence of sustainability commitments is the narrative contained in corporate reports. Misleading claims may pertain to both historical ESG performance and forward-looking pledges, often obscuring the issuer’s actual exposure to sustainability-related risks. As climate and environmental risks intensify, scrutiny around such claims is likely to increase, particularly where disclosures fail to demonstrate how corporate actions substantively contribute to global sustainability objectives. Against this backdrop, there is a growing need for transparent, scalable tools that can extract alert signals about greenwashing risk directly from firms’ sustainability narratives, especially in the green bond market where expectations around SDG alignment are high.

The academic literature has made important progress in conceptualizing greenwashing and ESG-washing and in documenting discrepancies between “talk” and “walk”, for example by comparing self-reported ESG disclosure to external ratings or environmental outcomes. However, empirical research on systematic *text-based* tools for issuing alert for potential greenwashing re-

¹See: https://www.oecd.org/en/publications/global-debt-report-2025_8ee42b13-en.html

mains comparatively scarce. Two gaps are particularly salient. First, existing measures rarely focus on the strategic omission of SDG-related information, despite the centrality of SDGs in the current sustainable finance architecture. Second, while several studies employ sentiment analysis, there is limited evidence on whether an opportunity-oriented environmental tone in sustainability reports can serve as a stand-alone alert indicator of greenwashing risk.

This study addresses these gaps by developing disclosure-based alert mechanisms for greenwashing in the context of European corporate green bond issuance. We propose an approach that combines Natural Language Processing (NLP) techniques with domain-specific transformer models based on BERT, including ClimateBERT (Bingler et al., 2022, 2024; Webersinke et al., 2022) and ESGBERT (Schimanski et al., 2024). In doing so, we rely exclusively on textual information extracted from sustainability reports, without imposing any ex-ante assumptions on actual environmental performance. A proprietary SDG-aligned dictionary is used to map targeted keywords to specific SDGs, leveraging the established alignment between the Green Bond Principles (GBPs) and the SDGs framework and building upon the prototype developed by CONSOB and the University of Trento (Paterlini et al., 2025).

Specifically, the study seeks to answer the following research questions. First, how can the omission of SDGs in sustainability disclosures be systematically analyzed and linked to factors such as type of sentence, sentiment, and other relevant textual characteristics, in order to develop potential greenwashing alerts in green bond issuances? Second, can the tone and sentiment of sustainability-related disclosures, quantified through domain-specific BERT-based models, be leveraged to develop alert metrics for greenwashing risk independently of SDG-related disclosure content?

This paper makes three key contributions to the literature on greenwashing, corporate social responsibility, and sustainable finance. First, it introduces the SDGs Omission Index (SDGOI), a metric designed to help investors, market participants, and regulators assess whether issuers effectively integrate their stated SDG commitments into communications and sustainability reports. By focusing on omission rather than solely on the presence of sustainability language, SDGOI provides a novel, disclosure-based lens on selective reporting practices in the green bond market. Second, the empirical analysis reveals a positive association between opportunity-oriented sentiment in environmental content and SDGOI, suggesting that tone can serve as a foundation for developing additional greenwashing alert metrics.

Third, building on this insight, the study develops and empirically validates the Environmental Sentiment Metric (ESM), a standardized measure of opportunity-oriented environmental language in sustainability narratives. We show that ESM scores are positively associated with ESG controversies (ESGCs) and greenwashing accusations (GWAs), even after controlling for ESG scores and firm size, underscoring their potential as reliable alert indicators of greenwashing risk in the green bond market.

The paper is organized as follows. Section 2 reviews the related literature. Section 3 presents the data. Section 4 introduces SDGOI, develops the hypotheses, and reports empirical results on SDGs omission. Section 5 then introduces the ESM and investigates its association with ESG controversies and greenwashing accusations. Section 6 concludes.

2 Literature Review

The increasing prevalence of sustainability-oriented corporate communication has intensified concerns regarding the authenticity of environmental claims, a phenomenon widely known as greenwashing (Delmas & Burbano, 2011). Despite the growing relevance of sustainability reporting, the empirical identification and quantification of greenwashing remain methodologically challenging and conceptually fragmented. The absence of a universally accepted definition, coupled with its multifaceted nature, has led to a proliferation of measurement approaches. Existing methods range from ESG disclosure–performance comparisons (e.g., Cho et al., 2012) to perception-driven proxies (e.g., Testa et al., 2018) and sentiment-based evaluations (e.g., Arena et al., 2015), highlighting the complexity of assessing the gap between symbolic and substantive sustainability actions (Lublóy et al., 2025).

Greenwashing is a relevant phenomenon not only in equity markets but also in corporate debt, given its implications for financing costs and investor trust (Peng & Xie, 2024; Roggi et al., 2024). A key research stream investigates the so-called green premium or greenium—that is, the additional cost (or discount) associated with environmentally friendly financial instruments. Empirical studies consistently document heterogeneity in the presence and magnitude of this effect, showing that investors may price both genuine sustainability and reputational risk (Baldi & Pandimiglio, 2022; Schmittmann & Gao, 2022; Xu et al., 2022). Within this debate, green bonds have attracted significant attention as emblematic instruments of sustainable finance. Ge

et al. (2025) examine the impact of green bond issuance on the ESG performance of Chinese listed firms, finding improvements in environmental and social dimensions. Issuance is thus interpreted as both a financing tool for sustainable projects and a reputational signal of corporate commitment. However, the study also uncovers evidence suggestive of greenwashing: improvements do not differ significantly across firms with varying pollution intensity or political ties, raising concerns about the authenticity of ESG enhancements. These findings align with broader critiques that disclosures, ratings, and voluntary statements, being largely self-reported, are prone to strategic manipulation (e.g., Cho et al., 2015; Christensen et al., 2021; Dyck et al., 2019).

A vast body of research emphasizes this disclosure–performance misalignment. Walker and Wan (2012) show that firms often engage in symbolic sustainability communication to protect legitimacy without substantive environmental change, ultimately undermining stakeholder trust and financial performance. Similarly, Yu et al. (2020) develop a peer-relative greenwashing score based on discrepancies between ESG disclosure and actual performance. Their findings reveal that large firms, despite extensive reporting, are particularly exposed to accusations of symbolic disclosure due to visibility and stakeholder pressure. Governance mechanisms, such as board independence and institutional ownership, can mitigate these risks, while external scrutiny, including foreign listings, reduces greenwashing incidence. More broadly, this literature conceptualizes greenwashing as a “talk” versus “walk” discrepancy (e.g., Kathan et al., 2025; Kim and Lyon, 2015; Lagasio, 2024), where symbolic communication exceeds substantive action.

Recent conceptual contributions expand the scope from “greenwashing” to “ESG-washing”. Todaro and Torelli (2024), for example, highlight how firms strategically appropriate ESG and circular economy terminology to construct reputational legitimacy in the absence of substantive alignment. This reinforces the argument that greenwashing is multidimensional, spanning environmental, social, and governance dimensions. At the same time, systematic reviews confirm both conceptual fragmentation and methodological diversity in this field. Lublóy et al. (2025) underline the reliance on perception-based measures and hypothetical cases, while Sneideriene and Legenzova (2025) use bibliometric mapping to document an increasing demand for rigorous, standardized metrics. Similarly, Ziolo et al. (2024) stress that reliance on theoretical frameworks such as stakeholder theory, legitimacy theory, and institutional theory is frequent but inconsistent, which complicates general-

ization and comparability.

Another important strand of literature critiques the reliability of ESG ratings. Since ratings rely heavily on self-reported disclosures, they are exposed to strategic manipulation. Moreover, methodological discrepancies among agencies further complicate their interpretation. Drempetic et al. (2020) show that firm size inflates ESG scores due to better disclosure capabilities rather than superior sustainability. This concern is reinforced by Kathan et al. (2025), who disentangle apparent from real environmental performance in European listed firms. While ESG scores correlate with narrative disclosure (apparent performance), they negatively associate with objective ecological outcomes, particularly among large firms. Treepongkaruna et al. (2024) reach similar conclusions in the U.S. context, showing that high ESG scores are not associated with lower carbon emissions, suggesting that ratings often function as reputational tools rather than performance indicators. These documented limitations of ESG scores have motivated the search for alternative, more granular measures of sustainability performance and misrepresentation.

To address these limitations, scholars increasingly turn to natural language processing (NLP) and sentiment-based approaches. These methods allow systematic analysis of textual data such as sustainability reports, press releases, earnings calls, and risk factor disclosures. By leveraging sentiment analysis, topic modeling, and semantic similarity, researchers can identify discrepancies between corporate narratives and operational outcomes. Early work on financial disclosure tone already highlighted the informativeness of sentiment measures. For instance, Koelbl (2019) document that the tone of MD&A disclosures conveys significant information for U.S. REIT performance, while Tsang et al. (2023) show that different sentiment types in risk factor disclosures matter for investors, particularly in the context of cybersecurity breaches. These insights from mainstream finance feed directly into sustainability studies, where textual sentiment is now applied to detect ESG-related misrepresentation.

Building on Loughran and McDonald (2016), recent sustainability literature uses sentiment dictionaries and machine learning tools to measure the tone, polarity, and readability of ESG communication. Studies such as Lucioni et al. (2020), Bingler et al. (2022), Bingler et al. (2024), and Moodaley and Telukdarie (2023) apply NLP-based metrics to detect divergences between corporate claims and actual performance, demonstrating their value as early warning signals. Conceptual frameworks such as that of Dorfleitner and Utz (2023) explicitly operationalize greenwashing as the gap between

“green talk” and “green walk,” while Lagasio (2024) propose the ESG-Washing Severity Index (ESGSI), which combines sentiment metrics with third-party ESG ratings to detect symbolic disclosure. Empirical evidence suggests that larger firms exhibit higher ESGSI scores, consistent with reputational incentives. Methodologically, this strand of work underscores the potential of advanced NLP techniques for scrutinizing sustainability reports and uncovering patterns of strategic communication.

Recent applications further highlight the strategic use of textual tone. Gorovaia and Makrominas (2024) show that U.S. firms subject to environmental penalties expand the length and positivity of CSR reports while reducing transparency, suggesting deliberate obfuscation. Similarly, bibliometric studies (e.g., Sneideriene and Legenzova, 2025) confirm that AI-based approaches are becoming a dominant frontier in detecting reputational manipulation. Importantly, these tools align with broader academic efforts to move from perception-based to verifiable, impact-oriented metrics. Yet, the scarcity of confirmed greenwashing cases remains a key challenge in validating proposed indicators.

Against this background, the present study contributes to the greenwashing and CSR literature in three ways. First, instead of comparing “talk” and “walk”, we focus exclusively on the “talk” dimension and examine whether specific linguistic patterns, particularly opportunity-oriented sentiment in environmentally themed sentences, can serve as alert signals for greenwashing risk. Second, we respond to concerns about the limitations of ESG ratings by developing granular, bottom-up measures derived directly from sustainability reports, rather than relying on aggregate scores that may be biased by firm size or methodological inconsistencies. Third, we situate this analysis within the specific context of the European green bond market, where SDG alignment is central yet under-examined from a textual perspective, thereby addressing a critical gap in attempts to capture greenwashing risk in sustainable debt instruments.

3 Data

3.1 Green Bond Data

This study focuses on the green bond market in Europe. We begin by clarifying the operational definition of a “green” bond. To ensure a robust clas-

sification, we rely on two independent data providers: LSEG and FactSet. LSEG applies the standards established by the International Capital Market Association (ICMA), whereas FactSet identifies green bonds by examining disclosures in the use-of-proceeds section of the bond prospectus or in the final terms and conditions of the issuance.

On the basis of this classification, issuers were ranked by the number of green bonds issued between 2013 and 2023. This starting point was chosen because prior to 2013 the issuance of green bonds was extremely limited, which renders it negligible for research purposes. Once issuers were sorted according to their issuance activity, their sustainability reports were retrieved through LSEG and supplemented with targeted manual web searches, restricted to the period 2019–2023. This temporal choice aligns with the evolution of the EU’s regulatory framework following the 2018 ‘Action Plan: Financing Sustainable Growth’². This initiative marked the beginning of a structured regulatory framework, making the subsequent period the most relevant for assessing disclosure practices and alignment with sustainability standards.

We exhaustively included all firms with seven or more green bond issuances. The remaining population was stratified into two groups (4–6 bonds and ≤ 3 bonds) and randomly sampled in proportion to their prevalence within the lower-issuance population. These proportions were applied exclusively during document retrieval; issuers without an accessible or English-language sustainability report were not replaced within their stratum. Although this method departs from the original issuance distribution—due to the complete inclusion of high-frequency issuers—it ensures broad market coverage, capturing approximately 70% of the total issued green-bond volume (as defined by LSEG and FactSet) in the period. This approach prioritizes representativeness in terms of issuance magnitude rather than issuer count and follows standard stratified sampling principles for skewed distributions (Tillé & Wilhelm, 2017).

The final dataset consists of 898 reports from 195 unique issuers, comprising 396 Sustainability Reports, 310 Other Non-Financial Reports, 77 Annual & Sustainability Reports, and 115 Annual Reports which also include sustainability information. For simplicity, the term “sustainability reports” will be used henceforth to refer to the documents analyzed. The distribution of

²See: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0097>

issuers’ domicile and sector by year is reported in Appendix 7.

3.2 NLP Methodology

The computation of text-based indicators relies on the NLP methodology presented in Paterlini et al. (2025). The procedure involves multiple stages.

Firstly, using ESGBERT (Schimanski et al., 2024), we identify phrases related to the environment in the sustainability report (label: *environmental* or *none*).

Secondly, ClimateBERT (Bingler et al., 2022, 2024; Webersinke et al., 2022) enables the identification of environmental claims in the report (label: *environmental claim* or *none*). An environmental claim is defined as a statement that expresses an opinion, a judgment, or verifiable information regarding climate or the environment. It differs from an environmental sentence in that the latter is a mere mention of environmental issues without expressing an opinion or verifiable information.

Thirdly, sentiment labels are assigned to sentences from the sustainability reports using ClimateBERT, classifying them based on their emphasis on risks (label: *risk*), on opportunities (label: *opportunity*), or a neutral tone (label: *neutral*).

Finally, SDG-related phrases are identified through a dictionary-based search in the sustainability report (i.e., *SDGsFound*) and compared with the declared SDGs (i.e., *SDGsDeclared*) of active green bonds in the reference year.

Our SDG dictionary follows the methodology described in Paterlini et al., 2025, where SDG-related keywords are constructed by mapping the Green Bond Principles to SDGs and integrating this mapping with the LSEG Use-of-Proceeds taxonomy. Once the dictionary with keywords associated for each SDG was defined, sustainability documents were searched to identify which SDG from the *a priori* declared set could be found in the reports (see also Paterlini et al., 2025). The NLP pipeline is deterministic and rule-based for SDG extraction, and transformer-based for environmental content, claims, and sentiment.

3.3 Variables

Based on the NLP-based methodology, we construct a set of text-based variables to quantify both the presence and the tone of content related to en-

environmental themes and the SDGs. Specifically, we compute the proportion of environmentally labeled sentences relative to the total number of sentences in each document (`EnvContent_Of_Total`), as well as the proportion of environmentally labeled claims (`EnvClaims_Of_Total`). In parallel, we calculate analogous metrics for SDG-related content, including the proportion of SDG-themed sentences (`SDGContent_Of_Total`) and the proportion of SDG-themed claims (`SDGClaims_Of_Total`).

To assess the tone of the content, we further compute the share of environmentally labeled sentences that are associated with opportunity-oriented or risky sentiment (`Sent_Opp_Of_EnvContent` and `Sent_Risk_Of_EnvContent`, respectively). Similarly, we derive the corresponding proportions for SDG-related sentences (`Sent_Opp_Of_SDGContent` and `Sent_Risk_Of_SDGContent`).

In addition to these text-based indicators, we incorporate a structural variable: the number of green bonds actively outstanding as of December 31 of the reporting year (`Active_Bonds`). While the relationship between the volume of outstanding bonds and the quality of sustainability disclosure has received limited attention in prior research, it is reasonable to expect that firms with a larger presence in the green bond market are subject to greater scrutiny by both investors and regulators. This increased visibility may, in turn, create stronger incentives for issuers to enhance the comprehensiveness, transparency, and reliability of their sustainability reporting. We refer to Table 6 in the Appendix for variables definition and Table 9 for the correlation matrix among variables.

4 SDGs Omission: Index, Hypotheses and Empirical Results

This section introduces the SDGs Omission Index (SDGOI), develops hypotheses on its relation to content and sentiment, and presents empirical evidence from a fixed-effects panel model.

4.1 SDGs Omission Index and Hypotheses

Sustainability disclosures are critical for evaluating companies' sustainability strategies and actions. Since the Paris Agreement in December 2015, corporate climate commitments have surged, increasing by nearly 25% between

2015 and 2020 (Bingler et al., 2022). Yet, questions remain as to whether these commitments reflect genuine environmental leadership or rather superficial responses to growing public scrutiny. Prior studies highlight the need to scrutinize the authenticity of corporate disclosures, warning of greenwashing risks when claims lack substantiation or measurable impact (e.g., Bingler et al., 2022; Schimanski et al., 2024).

This study examines weaknesses in corporate sustainability communication in the European green bond market, focusing on two elements: the selective omission of SDG-related information and the tone of sustainability disclosures. We first investigate whether textual sentiment and content characteristics are systematically linked to SDG omission, understood as a potential distinctive trait of greenwashing due to selective disclosure. The aim is to understand which types of communication are associated with a greater lack of SDG-related information, with a specific focus on “how businesses talk”.

By manipulating content and language in reports, firms can control the information provided, redirect public opinion and derive important competitive advantages. In the context of financial instruments, best practices outlined by market guidelines and standards such as the Green Bond Principles (GBPs) and the EU Taxonomy Regulation recommend that companies explicitly specify which SDGs the proceeds from these instruments aim to support. As such, the absence of clear references to the stated SDGs in corporate reporting to shareholders can reasonably be associated with a greenwashing alert, raising concerns about the transparency and authenticity of the company’s sustainability claims. Issuers might obscure the omission of information related to SDGs by inflating the text with general environmental statements lacking specific references or by emphasizing sentences with opportunity-oriented sentiment, while avoiding any mention of the climate risks associated with the projects financed by the green bond issuance.

This observation leads to our first two hypotheses:

Hypothesis 1 (Target information): *A greater use of SDG-themed sentences is associated with broader coverage of the SDGs declared by issuers.*

Hypothesis 2 (Dilution of information): *A greater use of environmental sentences is associated with less coverage of the SDGs declared by issuers.*

We therefore test the effects, on the omission of SDG information, of the concentration of SDG-themed phrases and environmental phrases. The interest of this analysis lies in the fact that companies with a high level of

cheap talk prioritize maintaining a positive public perception over making meaningful changes to their business practices (Bingler et al., 2024).

This research not only studies the content of the disclosure but also the way it is presented. In particular, the focus is on the sentiment of SDG- and environmental-themed sentences. While previous studies, such as Tsang et al. (2023) and Koelbl (2019), have shown that positive and negative sentiment types carry meaningful information, they do not extend their analysis to the specific context of green bonds. Here, we aim to test whether there is an association between coverage of SDG information and type of sentiment, distinguishing it as opportunity-oriented, risky, or neutral.

Therefore, we test two additional hypotheses:

Hypothesis 3 (SDG sentiment importance): *A higher presence of opportunity-oriented sentiment within SDG-related statements is associated with more coverage of the SDGs declared by the issuers.*

Hypothesis 4 (Environmental sentiment importance): *A higher presence of opportunity-oriented sentiment within environmental-related statements is associated with less coverage of the SDGs declared by the issuers.*

To operationalize SDG omission, we define the SDGs Omission Index (SDGOI) within the context of financial markets and following leading guidelines and regulatory frameworks such as the GBPs and the EU Taxonomy Regulation:

$$\text{SDGOI} = 1 - \frac{\text{SDGsFound}}{\text{SDGsDeclared}}, \quad (1)$$

where SDGsFound are the number of unique SDGs identified by the NLP approach within the sustainability report (see Section 3.2), while SDGsDeclared are the number of unique SDGs reported by LSEG with respect to the use of proceeds. The ratio indicates the proportion of declared SDGs that are actually identified in the reporting phase. Subtracting this ratio from one highlights the share of SDGs that remain unmentioned in the reports.

In the following analysis, only SDGs [1,2,3,6,7,8,9,11,12,13,14,15] are considered, as they map directly into GBPs. SDG-related sentences are extracted using the dictionary-based approach described in Section 3.2, which links keywords to each SDG by mapping the GBPs to the SDGs based on the ICMA SDG Mapping Report.³

³See: <https://www.icmagroup.org/assets/documents/Regulatory/Green-Bonds/>

4.2 Empirical Results on SDGs Omission

In order to test the hypotheses described above, we use a fixed-effects panel regression with heteroskedasticity-robust standard errors clustered at the firm level:

$$\begin{aligned} \text{SDGOI}_{i,t} = & \beta_0 + \beta_1 \text{EnvContent_Of_Total}_{i,t} + \beta_2 \text{EnvClaims_Of_Total}_{i,t} \\ & + \beta_3 \text{SDGContent_Of_Total}_{i,t} + \beta_4 \text{SDGClaims_Of_Total}_{i,t} \\ & + \beta_5 \text{Sent_Opp_Of_EnvContent}_{i,t} + \beta_6 \text{Sent_Risk_Of_EnvContent}_{i,t} \\ & + \beta_7 \text{SDG_Sent_Opp_Of_SDGContent}_{i,t} + \beta_8 \text{SDG_Sent_Risk_Of_SDGContent}_{i,t} \\ & + \beta_9 \text{Active_Bonds}_{i,t} + \epsilon_{i,t} \quad (2) \end{aligned}$$

for $i = 1, \dots, \text{Number of Issuers}$ and $t = 2019, \dots, 2023$, where we control for firm, year and sector fixed effects. Including sector fixed effects accounts for unobserved heterogeneity across entities that may influence the dependent variable but remain time-invariant. These controls are essential to mitigate omitted variable bias, as they absorb systematic differences between sectors that could otherwise confound the estimates (see Section 3 and Table 6 in the Appendix for variables description.)

Empirical results are shown in Table 1. The proportion of SDG sentences ($\text{SDGContent_Of_Total}$), as well as the proportion of SDG claims ($\text{SDGClaims_Of_Total}$), displays a strong and consistently negative relationship with the SDGOI across all model specifications ($p < 0.01$). This indicates, as expected, that a greater proportion of SDG content within a report is associated with more complete coverage of the declared SDGs. The negative coefficient suggests that integrating SDG content might serve as a mechanism for mitigating omission and selective disclosure, which is consistent with Hypothesis 1.

Conversely, Hypothesis 2 cannot be substantiated, as the share of environmental sentences ($\text{EnvContent_Of_Total}$) and the share of environmental claims ($\text{EnvClaims_Of_Total}$) are not statistically significant.

Hypotheses 3 and 4 cannot be rejected. The proportion of environmental sentences with opportunity-oriented sentiment ($\text{Sent_Opp_Of_EnvContent}$) exhibits a consistently significant positive relationship ($p < 0.1$ or $p < 0.05$) with the dependent variable across

Table 1: Fixed effects panel regression estimates of the impact of environmental and SDG-related content, claims and sentiment on SDGOI. Standard errors are heteroskedasticity-robust and clustered at the firm level. Year and sector fixed effects are included in all specifications. Robustness test including firm fixed effects w.r.t. also domicile and type of report provide qualitatively similar results (see Table 12 in Appendix). The linear fixed-effects (OLS) estimator is presented as the baseline specification because it delivers directly interpretable marginal effects and accommodates a rich set of fixed effects. By contrast, the correlated random-effects fractional logit estimator is used as a robustness check, as it ensures correct handling of the fractional nature of the dependent variable while preserving comparability with the linear FE structure. Findings on opportunity-oriented sentiment sentiment remain robust when estimating a fractional logit model with correlated random effects (Papke & Wooldridge, 2008); (see Table 13 in the Appendix).

	(1) SDGOI	(2) SDGOI	(3) SDGOI	(4) SDGOI
EnvContent_Of_Total	0.040 (0.119)			
EnvClaims_Of_Total		0.199 (0.375)		
SDGContent_Of_Total			-4.672*** (1.008)	
SDGClaims_Of_Total				-6.250*** (2.163)
Sent_Opp_Of_EnvContent	0.302** (0.144)	0.282* (0.150)	0.325** (0.128)	0.328** (0.132)
Sent_Risk_Of_EnvContent	-0.176 (0.311)	-0.166 (0.311)	-0.235 (0.295)	-0.191 (0.242)
SDG_Sent_Opp_Of_SDGContent	-0.426*** (0.066)	-0.427*** (0.066)	-0.412*** (0.065)	-0.410*** (0.035)
SDG_Sent_Risk_Of_SDGContent	-0.130 (0.132)	-0.125 (0.132)	-0.145 (0.133)	-0.148 (0.096)
Active_Bonds	0.001* (0.000)	0.001** (0.000)	0.001 (0.000)	0.001 (0.000)
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Sector FE	Yes	Yes	Yes	Yes
Observations	667	667	667	667
R^2	0.287	0.288	0.320	0.300
F-statistic	19.079***	19.103***	22.308***	20.233***

Note: Standard errors clustered at firm level in parentheses.

* p<0.1, ** p<0.05, *** p<0.01

all models. On the one hand, this finding aligns with the expectation that opportunity-oriented framing of environmental narratives is linked to SDG omission and potentially selective disclosure, a typical characteristic of greenwashing. On the other hand, opportunity-oriented sentiment of SDG sentences (`SDG_Sent_Opp_Of_SDGContent`) demonstrates a consistent and highly significant negative relationship ($p < 0.01$) with the SDGs Omission Index across all models. This implies that an opportunity-oriented tone in SDG-related communication is associated with broader disclosure of the declared SDGs.

Regarding risky sentiment, the coefficients of SDG-related sentences with risky sentiment (`SDG_Sent_Risk_Of_SDGContent`) as well as those of environmental sentences with risky sentiment (`Sent_Risk_Of_EnvContent`) are not significant.

Overall, the findings highlight the nuanced relationship between content, sentiment, and the strategic omission of SDG-related information in sustainability reports. These observations motivate the development of a complementary metric that captures opportunity-oriented environmental sentiment independently of SDG statements and assesses its association with ESG controversies and greenwashing accusations. A positive association would suggest its potential as an alert measure for greenwashing in a broader sense. The results are robust to a series of alternative specifications, including firm-level fixed effects, correlated random effects models, and winsorization of text-based variables, as detailed in Appendix 7.

5 A Sentiment-based Metric

The analysis of the SDGOI revealed a positive association with opportunity-oriented tone in corporate environmental statements. This finding indicates that textual information may serve as a useful indicator of discrepancies between the SDGs companies declare and those they actually pursue, thereby pointing to selective disclosure practices as a potential form of greenwashing. A limitation of this methodological approach is that it applies only to firms explicitly declaring a commitment to the SDGs. While this approach yields valuable insights, it inherently limits the sample size and the scope of applicability.

Building on previous findings that demonstrate a positive link between SDGOI and opportunity-oriented environmental sentiment, we address a

critical gap in sustainable finance by proposing a novel metric based on opportunity-oriented environmental sentiment, without being constrained by SDGs disclosures. By focusing on opportunity-oriented environmental sentiment, we aim to develop a greenwashing alert measure for enhancing transparency and risk assessment across the green bond market in Europe, thereby addressing a critical gap in both current literature and practice.

The Environmental Sentiment Metric

We introduce the Environmental Sentiment Metric (ESM), which is calculated as the standardized count of environment-related sentences exhibiting opportunity-oriented sentiment, following min-max normalization:

$$\text{ESM}_{i,t} = \frac{\text{Env_Sent_Opp}_{i,t} - \overline{\text{Env_Sent_Opp}_t}}{\sigma_{\text{Env_Sent_Opp}_t}} \quad (3)$$

where $\text{Env_Sent_Opp}_{i,t}$ is the number of environmental labeled sentences with opportunity-oriented sentiment for issuer i in year t after min-max normalization, $\overline{\text{Env_Sent_Opp}_t}$ is the mean of the number of environmental labeled sentences with opportunity-oriented sentiment across all issuers in year t , and $\sigma_{\text{Env_Sent_Opp}_t}$ is the corresponding standard deviation.

The ESM addresses critical challenges in textual analysis of sustainability disclosures by integrating min-max normalization and annual standardization. Min-max normalization accounts for stylistic differences (Sebastiani, 2002) and verbosity while preserving ordinality, avoiding the instability inherent in ratio-based metrics that conflate sentiment intensity with textual volume (Gentzkow & Shapiro, 2010; Loughran & McDonald, 2011). Standardization within each year ensures cross-sectional and temporal comparability, anchoring scores to evolving reporting norms (e.g., post-2015 SDGs adoption trends) and isolating firm-level deviations from industry-wide baselines (Hoberg & Phillips, 2016). By retaining absolute counts rather than ratios, the metric preserves the magnitude of opportunity-oriented sentences, distinguishing firms with substantively different disclosure scales—critical for greenwashing alerts. This methodology aligns with NLP best practices, mirroring TF-IDF weighting for corpus-adjusted relevance (Ramos et al., 2003) and Z-score standardization for robust feature engineering in document classification (Sebastiani, 2002).

5.1 ESG Controversies Score and Greenwashing Accusations

In the absence of actual greenwashing cases, empirical testing relies on issuers' ESG Controversy Scores (ESGCs) from LSEG and a self-developed Greenwashing Accusations (GWAs) metric. In line with prior literature, we define a controversy as any negative ESG-related event identified through media monitoring (Delmas and Burbano, 2011; Lyon and Maxwell, 2011), whereas an accusation refers to explicit claims of misleading or deceptive sustainability communication, whether in legal documents, regulatory actions, NGO reports, or investigative journalism. Furthermore, we also consider ESG Scores and the logarithm of the market capitalization as a proxy for issuer size, both provided by LSEG.

The ESG Score, as defined by LSEG, represents an overall company score based on the self-reported information in the environmental, social, and corporate governance pillars, with values from 0 (poorest ESG performance) to 100 (best grade). The ESGCs score, as defined by LSEG, is a measure of a company's exposure to environmental, social, and governance controversies and negative events related to sustainability issues which are tracked by global media. In the following analysis we construct a binary variable, still called ESGCs, which is equal to 0 if the firm had no controversies and 1 otherwise.

Figure 1 in the top panel reports the relationship between ESG Scores and the presence of ESG controversies in the sample period. The boxplots suggest that firms facing controversies tend to exhibit higher ESG Scores. This result is in line with the literature (e.g., Kathan et al., 2025; Kim and Lyon, 2015; Marquis et al., 2016) and could be explained by considering that firms with better scores receive more media and stakeholder attention and, as a consequence, could be more likely subject to disputes.

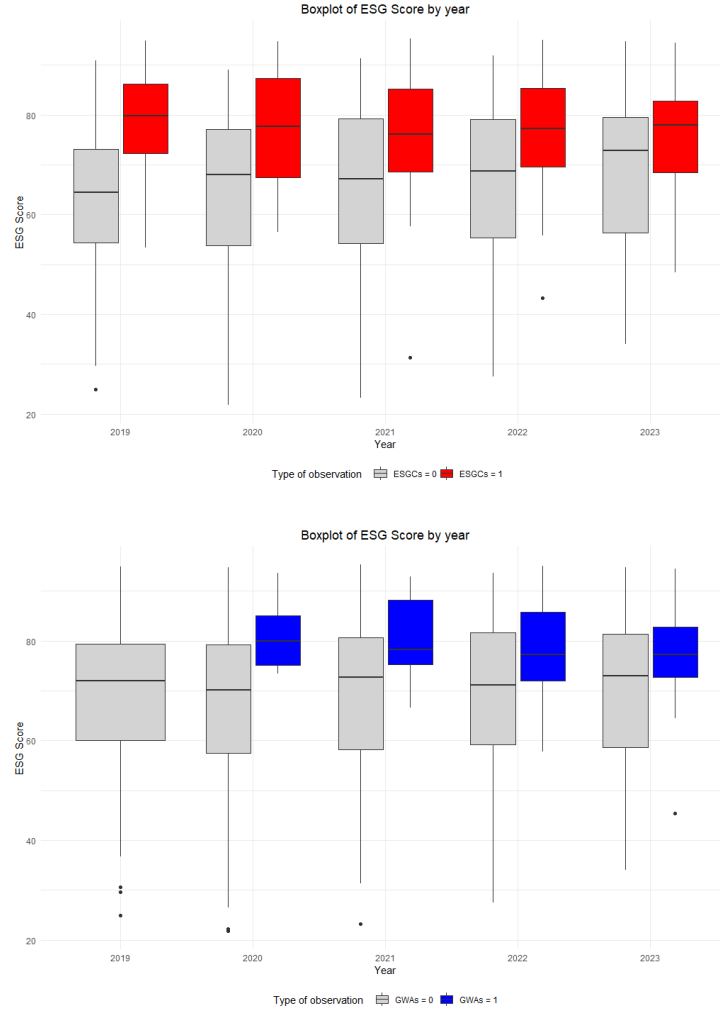


Figure 1: Boxplots of ESG Scores. Top panel: issuers with ESGCs = 0 (gray) and ESGCs = 1 (red). Bottom panel: issuers with GWAs = 0 (gray) and GWAs = 1 (blue).

We also introduce the Greenwashing Accusations (GWAs) variable, defined as a binary indicator that identifies whether a firm was subject to formal legal allegations or public accusations of greenwashing between 2019 and 2023. To account for potential late disclosures in sustainability reports, the coverage of 2023 accusations was extended through March 2024.

Following established literature on corporate environmental misconduct

(Delmas & Burbano, 2011; Lyon & Montgomery, 2015), the variable includes both formal legal actions (e.g., regulatory sanctions or lawsuits) and non-legal but publicly documented allegations (e.g., NGO reports or media investigations), recognizing that reputational damage can arise even without litigation (Seele & Gatti, 2017).

We constructed the variable through a two-phase methodology, integrating AI-assisted data collection with manual verification. Initial data gathering employed large language models (i.e., DeepSeek and ChatGPT), which were systematically prompted to identify legal and non-legal allegations against each target company between January 2019 and March 2024. To ensure traceability, each query explicitly requested source citations and verifiable URLs. As recommended in recent LLM auditing research (Tonmoy et al., 2024), all AI-collected outputs were subject to strict manual source verification to avoid hallucinations and ensure provenance. Rigorous manual validation was conducted, involving both direct examination of the AI-provided sources and supplementary searches using identical parameters on Google. This dual approach, leveraging computational efficiency while maintaining human oversight, was implemented to mitigate potential hallucinations or omissions inherent in automated retrieval systems, thereby enhancing the dataset’s reliability.

While this approach provides a useful measure of greenwashing risk, it may underreport accusations against private firms (Kim & Lyon, 2015). Nevertheless, the variable aligns with reputational risk theory (Walker & Wan, 2012), capturing both regulatory and societal dimensions of greenwashing, and is particularly relevant for green bond issuers given the region’s stringent disclosure standards (EU Taxonomy Regulation, 2020).

The boxplots in the bottom panel of Figure 1 highlight a positive relationship also between ESG Scores and GWAs. Both ESGCs and GWAs, as a consequence, appear to be coherent indicators of reputational risk exposure, suggesting that they capture overlapping dimensions of ESG-related scrutiny, albeit with differing scope and intensity. GWA concentration among high-ESG firms mirrors the broader controversy trend, reinforcing the idea that superior ESG performance attracts heightened attention by stakeholders that may uncover both generalized controversies and targeted allegations. Both ESGCs and GWAs serve as indicators of potential issues related to greenwashing among green bond issuers. ESGCs typically cover a broad range of incidents and apply to a larger set of companies, with 208 reported cases. In contrast, GWAs (coded as 1) are more narrowly defined and pertain to

a smaller subset of issuers (50 in total), often highlighting specific cases in which firms are suspected of misleadingly portraying their sustainability efforts. As expected, there is strong overlap between the two, with 34 out of the 50 GWAs also captured within ESGCs.

Furthermore, if we consider the distribution of firm size by ESGCs and GWAs, we can notice a positive dependence between those variables. These findings are consistent with prior literature which suggests that larger firms have broader operations and higher ESG values, and that greater visibility and disclosure by large firms increases the probability of detection and reporting of controversies (e.g., Gregory, 2024; Kathan et al., 2025; Lins et al., 2017).

5.2 Empirical Analysis

In the following, we conduct both univariate and multivariate analyses to empirically test ESM.

5.2.1 Univariate Analysis

We begin with a univariate analysis to compare group means between firms with ESGCs or GWAs equal to 1 and those equal to 0, in order to test whether the two groups differ significantly across key variables. To assess statistical significance, we applied two tests: an independent samples t-test, which evaluates whether the means of each variable differ significantly between groups, and a Wilcoxon rank-sum test, which assesses differences in medians.

Table 2 presents group differences between observations with and without ESG controversies (ESGCs equal to 1 vs ESGCs equal to 0) at time t , regarding ESM, ESG Scores, the difference in ESG scores from time $t + 1$ to t (i.e. Δ ESG Score), and Size at time t .

The analysis reveals statistically significant differences between firms with and without ESGCs. Results related to ESM show that being a firm with controversies tends to be associated with a higher frequency of opportunity-oriented environmental sentences in sustainability reports. On the basis of this evidence, firms employing assertive language are, in fact, those most frequently subject to controversy. This result is in line with one of our research questions: opportunity-oriented sentiment could be used to build a greenwashing risk alert indicator.

As expected from inspecting also Figure 1, firms under dispute exhibit significantly higher ESG Scores both in t and $t + 1$.

Several strands of academic research have sought to explain the counterintuitive finding that firms with higher ESG scores may also be more frequently associated with ESG-related controversies (e.g., Kathan et al., 2025; Kim and Lyon, 2015). First, firms with elevated ESG ratings are often large, publicly traded entities that actively promote their sustainability credentials. This heightened visibility could make them more susceptible to scrutiny from stakeholders, media outlets, and regulatory bodies. Consequently, any deviation from stated ESG commitments is more likely to be detected and publicized, resulting in a greater incidence of reported controversies. Furthermore, firms with robust ESG programs often engage in multifaceted initiatives across diverse geographies and sectors. The complexity inherent in managing such programs increases the likelihood of operational lapses or ethical breaches.

Lastly, it is important to underline the temporal granularity of ESG scores provided by LSEG. ESG performance is typically reported on an annual basis, reflecting year-end values across a range of environmental, social, and governance indicators. This annual frequency implies that ESG scores may not fully capture intra-year developments, including short-term controversies or corrective actions taken by firms in response to stakeholder pressure. The lagged nature of ESG data can contribute to a disconnect between a firm’s current ESG performance and its most recently published ESG score. For instance, a firm may receive a high ESG rating based on the previous year’s disclosures, while simultaneously facing reputational challenges or controversies that emerge in the current reporting period. This temporal misalignment may partially explain the observed positive correlation between ESG scores and ESG controversies in empirical studies: ESG scores provided by LSEG may incorporate information regarding controversies over time. Indeed, the improvement in ESG Score (Δ ESG Score) is less pronounced for controversial firms (0.77) compared to non-controversial ones (2.07). While ESG Scores tend to improve over time, those of controversial firms improve less, suggesting that the score incorporates information regarding disputes only with some delay.

The same group comparison analysis was replicated using GWAs as the classification variable, yielding qualitatively consistent results, which are shown in Table 3. The direction and statistical significance of key findings—including higher opportunity-oriented sentiment, larger size, elevated

ESG scores, and attenuated post-event ESG improvement—closely mirrored those observed for ESGCs. This coherence across the two related measures strengthens the inference that both variables capture underlying firm-level traits associated with reputational risk and strategic disclosure behavior. Such consistency supports the use of either variable—ESGC or GWA—or both complementarily in analyses of ESG-related disclosure strategies. This is particularly relevant in the multivariate analysis, where we consider their relationship with ESM, ESG Scores, and firm size.

Table 2: Univariate analysis results comparing firms with ESGCs equal to 1 or 0. Differences in means and medians for the ESM and ESG Scores are evaluated using t-tests and Wilcoxon rank-sum tests.

Variable	Mean (ESGCs=0)	Mean (ESGCs=1)	Mean Diff.	t-test	Wilcoxon
ESM _t	-0.13	0.26	0.39	***	***
ESM _{t+1}	-0.13	0.32	0.45	***	***
ESG Score _t	65.88	76.90	11.02	***	***
ESG Score _{t+1}	67.28	77.82	10.54	***	***
ΔESG Score	2.07	0.77	-1.30	**	**
Size _t	23.00	24.10	1.10	***	***

Note: * p<0.1, ** p<0.05, *** p<0.01

Table 3: Univariate analysis results comparing firms with GWAs equal to 1 or 0. Differences in means and medians for the ESM and ESG Scores are evaluated using t-tests and Wilcoxon rank-sum tests.

Variable	Mean (GWAs=0)	Mean (GWAs=1)	Mean Diff.	t-test	Wilcoxon
ESM _t	-0.04	0.42	0.46	**	***
ESM _{t+1}	0.00	0.45	0.45	*	***
ESG Score _t	68.80	78.29	9.49	***	***
ESG Score _{t+1}	70.41	80.60	10.19	***	***
ΔESG Score	1.69	0.37	-1.32		
Size _t	23.30	24.20	0.90	***	***

Note: * p<0.1, ** p<0.05, *** p<0.01

5.2.2 Multivariate Analysis

Building on the univariate analysis, we estimate two separate logistic regression models to examine how key explanatory variables affect the likelihood of ESGCs and GWAs. Each model includes fixed effects for year and sector to control for unobserved heterogeneity and to isolate the effect of the independent variables on the outcomes of interest.

In particular, we estimate:

$$\text{logit}(Pr[Y_{i,t}]) = \beta_0 + \beta_1 \text{ESM}_{i,t} + \beta_2 \text{ESG Score}_{i,t} + \beta_3 \text{Size}_{i,t} + FE_t, \quad (4)$$

$$\text{logit}(Pr[Y_{i,t+1}]) = \beta_0 + \beta_1 \text{ESM}_{i,t} + \beta_2 \text{ESG Score}_{i,t} + \beta_3 \text{Size}_{i,t} + FE_t, \quad (5)$$

where Y is either ESGCs or GWAs, and FE_t refers to year and sector fixed effects.

Estimates with ESGCs and GWAs at time t are reported in Table 4, while Table 5 shows the results when considering ESGCs and GWAs at time $t + 1$.

Counterintuitively, higher ESG Scores show a positive association both with ESGCs at time t and $t+1$, and with GWAs at time t and $t+1$, potentially reflecting “ESG overclaiming” (Lyon & Montgomery, 2015) or measurement limitations in capturing genuine sustainability efforts, as noted in critiques of ESG score divergence (Berg et al., 2022).

The analysis identifies that ESM is persistently linked with ESGCs both contemporaneously and with a one-period lag. ESM is also positively linked with GWAs within the same year when the ESG Score is not explicitly accounted for. The positive association aligns with theoretical frameworks on opportunity-oriented environmental disclosure, where firms engage in symbolic sustainability practices to manipulate stakeholder perceptions while avoiding substantive environmental commitments (Lyon & Montgomery, 2015; Testa et al., 2018).

When ESM and ESG Score are included jointly in the full models, the results confirm that both predictors maintain their statistical significance in explaining ESGCs. This suggests that ESM provides information not fully subsumed by ESG scores, consistent with prior research showing that sentiment-based indicators capture dimensions of market perception distinct from rating-based measures (Loughran & McDonald, 2016).

The attenuation of ESM’s significance in some full specifications is not related to multicollinearity issues (see correlation and VIF diagnostics in the Appendix), but rather is consistent with the partial conceptual overlap between sentiment-based and score-based ESG proxies. This dynamic has been documented in related work showing that alternative sustainability metrics often share explanatory content but contribute complementary information depending on the outcome of interest (Chatterji et al., 2016; Dorfleitner et al., 2015).

Firm size emerges as a positive and highly significant variable across nearly all specifications. Larger firms have a higher probability of facing both ESG controversies and greenwashing accusations, both in the current and subsequent year. This finding provides robust support for the “visibility hypothesis,” which posits that larger firms attract greater scrutiny from media, activists, and regulators, making them more susceptible to public controversies (Brammer et al., 2006; Clarkson et al., 2008).

The results underscore the importance of distinguishing between authentic environmental stewardship and strategic posturing, as standardized sentiment metrics like ESM can act as alert indicators of latent legitimacy crises.

6 Conclusion

Green bonds have emerged as a cornerstone of sustainable finance, enabling companies and governments to raise capital for environmentally beneficial projects. However, the rapid growth of this market has raised concerns about the credibility and verifiability of environmental claims, making it essential to develop robust tools for issuing potential alerts on greenwashing.

Against this backdrop, this study introduces the Sustainable Development Goals Omission Index (SDGOI), demonstrating that the omission of SDG-related information in sustainability reporting is systematically associated with an opportunity-oriented environmental tone. Empirical evidence reveals that such omissions reflect consistent patterns of selective disclosure, a practice commonly linked to greenwashing, thus providing a transparent and interpretable signal of misalignment between declared and reported SDGs. The SDGOI therefore emerges as a useful tool for detecting disclosure gaps that mirror opportunity-oriented behavior, offering researchers and practitioners a new lens to understand, monitor, and potentially capture greenwashing practices in sustainability communication.

Table 4: Logistic regression assessing the effect of ESM, ESG Scores and Size on the probability of ESG controversies (ESGCs) and greenwashing accusations (GWAs) at time t . The models control for year and sector fixed effects and account for clustering at the issuer level. Robustness checks confirm that results are qualitatively similar after including total sentence counts or its logarithmic transformation as control variable. Results are available upon request.

Dependent Variable:	ESGCs _t			GWAs _t		
<i>Variables</i>						
ESM _t	0.600*** (0.202)		0.347** (0.165)	0.327** (0.129)		0.179 (0.134)
ESG Score _t		0.065*** (0.012)	0.059*** (0.012)		0.057*** (0.014)	0.053*** (0.014)
Size _t	0.463*** (0.117)	0.368*** (0.112)	0.367*** (0.112)	0.309*** (0.113)	0.271** (0.116)	0.267** (0.118)
<i>Fixed-effects</i>						
Year	Yes	Yes	Yes	Yes	Yes	Yes
Sector	Yes	Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>						
Observations	623	623	623	502	502	502
Squared Correlation	0.228	0.277	0.289	0.110	0.127	0.129
Pseudo R2	0.156	0.203	0.211	0.069	0.103	0.100

Clustered (IssuerID) standard errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 5: Logistic regression assessing the effect of ESM, ESG Scores and Size on the probability of ESG controversies (ESGCs) and greenwashing accusations (GWAs) at time $t + 1$. The models control for year and sector fixed effects and account for clustering at the issuer level. Robustness checks confirm that these associations are not driven by document length; coefficients remain stable after including total sentence counts and their logarithmic transformation as controls.

Dependent Variable:	ESGCs _{t+1}			GWAs _{t+1}		
<i>Variables</i>						
ESM _t	0.604*** (0.192)		0.362** (0.154)	0.086 (0.201)		-0.135 (0.205)
ESG Score _t		0.059*** (0.012)	0.053*** (0.012)		0.060*** (0.014)	0.063*** (0.015)
Size _t	0.396*** (0.123)	0.300*** (0.114)	0.301*** (0.114)	0.262** (0.107)	0.202* (0.115)	0.203* (0.115)
<i>Fixed-effects</i>						
Year	Yes	Yes	Yes	Yes	Yes	Yes
Sector	Yes	Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>						
Observations	482	482	482	464	464	464
Squared Correlation	0.205	0.258	0.270	0.092	0.128	0.130
Pseudo R2	0.126	0.167	0.176	0.037	0.093	0.088

Clustered (IssuerID) standard errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Building on this, we extend the analysis beyond issuers that explicitly declare SDGs and introduce ESM. This metric quantifies the extent to which environmental statements in corporate sustainability reports adopt an opportunity-oriented tone, thereby providing a broader perspective on rhetorical strategies associated with greenwashing. Univariate and multivariate analyses consistently show that high ESM values are positively associated with ESG controversies in both the current and subsequent year, as well as with greenwashing accusations in the same year. These results suggest that opportunity-oriented sentiment in environmental narratives is not merely a stylistic choice, but a potential marker of credibility risks in sustainability disclosure.

In this sense, ESM underscores how sentiment-driven rhetoric contributes to the construction of sustainability narratives that may overstate commitment and underplay challenges, reinforcing the role of language as a strategic tool in shaping stakeholder perceptions.

Looking ahead, this line of research offers several promising paths. First, refining SDGOI and ESM through AI and machine learning techniques could improve their contextual sensitivity and enhance the detection of omission patterns and sentiment-driven misrepresentation. Second, applying these metrics to a broader and more diverse pool of issuers would strengthen their generalizability and robustness. A wider availability of documented greenwashing cases would be especially valuable for validating the predictive power of these indicators. Future research should also examine the adaptability of the proposed model to evolving regulatory frameworks, assessing its ability to remain effective as sustainability disclosure requirements are revised and expanded. Finally, embedding SDGOI and ESM within the evolving infrastructure of sustainable finance, whether in disclosure standards, independent assurance practices or ESG data platforms, could significantly enhance their practical utility. Such integration would not only provide investors and regulators with reliable tools for evaluating the credibility of sustainability disclosures but also create stronger incentives for issuers to engage in truthful, transparent, and accountable reporting. In this way, SDGOI and ESM may serve not only as diagnostic instruments for greenwashing alerts but also as catalysts for a cultural shift toward greater accountability and integrity in sustainability communication.

References

- Allison, P. D. (2012). *Logistic regression using SAS: Theory and application*. SAS institute.
- Arena, C., Bozzolan, S., & Michelon, G. (2015). Environmental reporting: Transparency to stakeholders or stakeholder manipulation? an analysis of disclosure tone and the role of the board of directors. *Corporate Social Responsibility and Environmental Management*, 22(6), 346–361.
- Baldi, F., & Pandimiglio, A. (2022). The role of ESG scoring and greenwashing risk in explaining the yields of green bonds: A conceptual framework and an econometric analysis. *Global Finance Journal*, 52, 100711.
- Berg, F., Kölbel, J. F., & Rigobon, R. (2022). Aggregate confusion: The divergence of esg ratings. *Review of Finance*, 26(6), 1315–1344.
- Bingler, J. A., Kraus, M., Leippold, M., & Webersinke, N. (2022). Cheap talk and cherry-picking: What climatebert has to say on corporate climate risk disclosures. *Finance Research Letters*, 47, 102776.
- Bingler, J. A., Kraus, M., Leippold, M., & Webersinke, N. (2024). How cheap talk in climate disclosures relates to climate initiatives, corporate emissions, and reputation risk. *Journal of Banking & Finance*, 164, 107191.
- Brammer, S., Brooks, C., & Pavelin, S. (2006). Corporate social performance and stock returns: Uk evidence from disaggregate measures. *Financial Management*, 35(3), 97–116.
- 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.
- Cho, C. H., Guidry, R. P., Hageman, A. M., & Patten, D. M. (2012). Do actions speak louder than words? an empirical investigation of corporate environmental reputation. *Accounting, Organizations and Society*, 37(1), 14–25.
- Cho, C. H., Laine, M., Roberts, R. W., & Rodrigue, M. (2015). Organized hypocrisy, organizational façades, and sustainability reporting. *Accounting, Organizations and Society*, 40, 78–94.
- Christensen, H. B., Hail, L., & Leuz, C. (2021). Mandatory csr and sustainability reporting: Economic analysis and literature review. *Review of Accounting Studies*, 26(3), 1176–1248.

- Clarkson, P. M., Li, Y., Richardson, G. D., & Vasvari, F. P. (2008). Revisiting the relation between environmental performance and environmental disclosure: An empirical analysis. *Accounting, Organizations and Society*, 33(4-5), 303–327.
- Delmas, M. A., & Burbano, V. C. (2011). The drivers of greenwashing. *California Management Review*, 54(1), 64–87.
- Dorfleitner, G., Halbritter, G., & Nguyen, M. (2015). Measuring the level and risk of corporate responsibility—an empirical comparison of different esg rating approaches. *Journal of Asset Management*, 16(7), 450–466.
- Dorfleitner, G., & Utz, S. (2023). Green, green, it’s green they say: A conceptual framework for measuring greenwashing on firm level. *Review of Managerial Science*, 1–24.
- 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*, 333–360.
- Dyck, A., Lins, K. V., Roth, L., & Wagner, H. F. (2019). Do institutional investors drive corporate social responsibility? international evidence. *Journal of Financial Economics*, 131(3), 693–714.
- Ge, P., Liu, Y., Tang, C., & Zhu, R. (2025). Green bonds and corporate environmental social and governance performance: Innovative approaches to identifying greenwashing in green bond markets. *Corporate Social Responsibility and Environmental Management*, 32(1), 1060–1078.
- Gentzkow, M., & Shapiro, J. M. (2010). What drives media slant? evidence from us daily newspapers. *Econometrica*, 78(1), 35–71.
- Gorovaia, N., & Makrominas, M. (2024). Identifying greenwashing in corporate-social responsibility reports using natural-language processing. *European Financial Management*.
- Gregory, R. P. (2024). How greenwashing affects firm risk: An international perspective. *Journal of Risk and Financial Management*, 17(11), 526.
- Hadi, A. S. (1992). A new measure of overall potential influence in linear regression. *Computational Statistics & Data Analysis*, 14(1), 1–27.
- Hoberg, G., & Phillips, G. (2016). Text-based network industries and endogenous product differentiation. *Journal of Political Economy*, 124(5), 1423–1465.
- Huber, P. J. (1981). *Robust statistics*. Wiley.
- Kathan, M. C., Utz, S., Dorfleitner, G., Eckberg, J., & Chmel, L. (2025). What you see is not what you get: Esg scores and greenwashing risk. *Finance Research Letters*, 74, 106710.

- Kim, E.-H., & Lyon, T. P. (2015). Greenwash vs. brownwash: Exaggeration and undue modesty in corporate sustainability disclosure. *Organization Science*, 26(3), 705–723.
- Koelbl, M. (2019). *MD&A disclosure and performance of US REITs: The information content of textual tone* (tech. rep.). European Real Estate Society (ERES).
- Lagasio, V. (2024). Esg-washing detection in corporate sustainability reports. *International Review of Financial Analysis*, 96, 103742.
- Lins, K. V., Servaes, H., & Tamayo, A. (2017). Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis. *The Journal of Finance*, 72(4), 1785–1824.
- Loughran, T., & McDonald, B. (2011). When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *The Journal of Finance*, 66(1), 35–65.
- Loughran, T., & McDonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, 54(4), 1187–1230.
- Lublóy, Á., Keresztúri, J. L., & Berlinger, E. (2025). Quantifying firm-level greenwashing: A systematic literature review. *Journal of Environmental Management*, 373, 123399.
- Luccioni, A., Baylor, E., & Duchene, N. (2020). Analyzing sustainability reports using natural language processing. *arXiv preprint arXiv:2011.08073*.
- Lyon, T. P., & Maxwell, J. W. (2011). Greenwash: Corporate environmental disclosure under threat of audit. *Journal of Economics & Management Strategy*, 20(1), 3–41.
- Lyon, T. P., & Montgomery, A. W. (2015). The means and end of greenwash. *Organization & Environment*, 28(2), 223–249.
- Marquis, C., Toffel, M. W., & Zhou, Y. (2016). Scrutiny, norms, and selective disclosure: A global study of greenwashing. *Organization Science*, 27(2), 483–504.
- Menard, S. (2001). *Applied logistic regression analysis*. SAGE publications.
- Moodaley, W., & Telukdarie, A. (2023). Greenwashing, sustainability reporting, and artificial intelligence: A systematic literature review. *Sustainability*, 15(2), 1481.
- Papke, L. E., & Wooldridge, J. M. (2008). Panel data methods for fractional response variables with an application to test pass rates. *Journal of econometrics*, 145(1-2), 121–133.

- Paterlini, S., Nicolodi, A., Gentile, M., Manzillo, V., Sancilio, M., & Deriu, P. (2025). Greenwashing Alert System for EU Green Bonds. The CONSOB-University of Trento Prototype. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.5379964>
- Peng, Q., & Xie, Y. (2024). Esg greenwashing and corporate debt financing costs. *Finance Research Letters*, 69, 106012.
- Ramos, J., et al. (2003). Using tf-idf to determine word relevance in document queries. *Proceedings of the First Instructional Conference on Machine Learning*, 242(1), 29–48.
- Roggi, O., Bellardini, L., & Conticelli, S. (2024). Effects of esg performance and sustainability disclosure on gss bonds' yields and spreads: A global analysis. *Finance Research Letters*, 68, 105988.
- Rousseeuw, P. J., & Leroy, A. M. (2003). *Robust regression and outlier detection*. John Wiley & sons.
- Schimanski, T., Reding, A., Reding, N., Bingler, J., Kraus, M., & Leippold, M. (2024). Bridging the gap in esg measurement: Using NLP to quantify environmental, social, and governance communication. *Finance Research Letters*, 61, 104979.
- Schmittmann, J. M., & Gao, Y. (2022). *Green bond pricing and greenwashing under asymmetric information*. International Monetary Fund.
- Sebastiani, F. (2002). Machine learning in automated text categorization. *ACM Computing Surveys (CSUR)*, 34(1), 1–47.
- Seele, P., & Gatti, L. (2017). Greenwashing revisited: In search of a typology and accusation-based definition incorporating legitimacy strategies. *Business Strategy and The Environment*, 26(2), 239–252.
- Sneideriene, A., & Legenzova, R. (2025). Greenwashing prevention in environmental, social, and governance (esg) disclosures: A bibliometric analysis. *Research in International Business and Finance*, 74, 102720.
- Testa, F., Boiral, O., & Iraldo, F. (2018). Internalization of environmental practices and institutional complexity: Can stakeholders pressures encourage greenwashing? *Journal of Business Ethics*, 147, 287–307.
- Tillé, Y., & Wilhelm, M. (2017). Probability sampling designs: Principles for choice of design and balancing. *Statistical Science*, 176–189.
- Todaro, D. L., & Torelli, R. (2024). From greenwashing to esg-washing: A focus on the circular economy field. *Corporate Social Responsibility and Environmental Management*, 31(5), 4034–4046.

- Tonmoy, S., Zaman, S., Jain, V., Rani, A., Rawte, V., Chadha, A., & Das, A. (2024). A comprehensive survey of hallucination mitigation techniques in large language models. *arXiv preprint arXiv:2401.01313*, 6.
- Treepongkaruna, S., Au Yong, H. H., Thomsen, S., & Kyaw, K. (2024). Greenwashing, carbon emission, and ESG. *Business Strategy and The Environment*, 33(8), 8526–8539.
- Tsang, R. C., Baldwin, A. A., Hair, J. F., Affuso, E., & Lahtinen, K. D. (2023). The informativeness of sentiment types in risk factor disclosures: Evidence from firms with cybersecurity breaches. *Journal of Information Systems*, 37(3), 157–190.
- Walker, K., & Wan, F. (2012). The harm of symbolic actions and greenwashing: Corporate actions and communications on environmental performance and their financial implications. *Journal of Business Ethics*, 109, 227–242.
- Webersinke, N., Kraus, M., Bingler, J., & Leippold, M. (2022). Climate-BERT: A Pretrained Language Model for Climate-Related Text. *Proceedings of AAAI 2022 Fall Symposium: The Role of AI in Responding to Climate Challenges*. <https://doi.org/https://doi.org/10.48550/arXiv.2212.13631>
- Xu, G., Lu, N., & Tong, Y. (2022). Greenwashing and credit spread: Evidence from the chinese green bond market. *Finance Research Letters*, 48, 102927.
- Yu, E. P.-y., Van Luu, B., & Chen, C. H. (2020). Greenwashing in environmental, social and governance disclosures. *Research in International Business and Finance*, 52, 101192.
- Zioło, M., Bąk, I., & Spoz, A. (2024). Literature review of greenwashing research: State of the art. *Corporate Social Responsibility and Environmental Management*, 31(6), 5343–5356.

7 Appendix

Table 6: Variables definitions for the SDGOI fixed-effects panel model

Variable	Definition
SDGOI	$\text{SDGOI} = \frac{\text{SDGsFound}}{\text{SDGsDeclared} + \text{SDGsIdentified}}$ Sustainable Development Goals Omission Index, defined as $1 - \frac{\text{SDGsFound}}{\text{SDGsDeclared} + \text{SDGsIdentified}}$, where SDGsFound are the distinct number of the 12 relevant SDGs identified by the NLP approach within the sustainability report and SDGsDeclared are the distinct number of the 12 SDGs reported by LSEG with respect to the use of proceeds.
EnvContent_Of_Total	Share of sentences in the report that are labeled as environmental (by ESG-BERT) over the total number of sentences in the document.
EnvClaims_Of_Total	Share of sentences labeled as environmental claims (by ClimateBERT) over the total number of sentences in the document.
SDGContent_Of_Total	Share of sentences containing SDG-related content (identified via the SDG dictionary) over the total number of sentences in the document.
SDGClaims_Of_Total	Share of sentences classified as SDG-related claims over the total number of sentences in the document.
Sent_Opp_Of_EnvContent	Proportion of environmental sentences whose sentiment is labeled as Opportunity (by ClimateBERT), relative to all environmental sentences.
Sent_Risk_Of_EnvContent	Proportion of environmental sentences whose sentiment is labeled as Risk, relative to all environmental sentences.
SDG_Sent_Opp_Of_SDGContent	Proportion of SDG-related sentences whose sentiment is labeled as Opportunity, relative to all SDG-related sentences.
SDG_Sent_Risk_Of_SDGContent	Proportion of SDG-related sentences whose sentiment is labeled as Risk, relative to all SDG-related sentences.
Active_Bonds	Number of green bonds issued by the firm that are actively outstanding as of 31 December of year t .

Descriptive Statistics - Full sample.

Table 7: Sector Distribution per Year

Sector	2019	2020	2021	2022	2023
Agency	1	1	1	1	1
Banks	48	52	56	57	54
Electric Power	18	21	23	22	21
Energy Company	4	4	4	4	4
Gas Distribution	2	2	2	2	2
Manufacturing	36	41	42	42	39
Other Financial	26	34	34	35	30
Service Company	15	18	19	19	19
Communication	4	4	5	5	5
Transportation	2	2	4	5	5
Total	156	179	190	192	180

Note: Green bonds, which belong to the ‘Agency’ sector on the basis of LSEG, are classified as corporate bonds by FactSet and for this reason they are included.

Table 8: Domicile distribution per Year

Domicile	2019	2020	2021	2022	2023
AT	5	6	7	7	7
BE	3	4	5	5	4
CZ	0	0	0	1	1
DE	29	29	31	30	29
DK	8	8	8	8	8
ES	11	12	12	13	12
FI	7	7	7	7	7
FR	19	24	28	27	27
GR	5	5	5	5	3
HU	1	1	1	1	1
IE	2	5	5	5	4
IT	18	20	20	20	18
LU	3	4	4	4	4
NL	12	17	16	18	17
PL	3	4	4	4	3
PT	1	1	1	1	0
SE	29	31	35	35	34
SI	0	1	1	1	1
Total	156	179	190	192	180

Table 9: Correlation between variables (SDGOI). Variables: (1) EnvContent_Of_Total = share of environmental content in total text; (2) EnvClaims_Of_Total = share of environmental claims in total text; (3) SDGContent_Of_Total = share of SDG content in total text; (4) SDGClaims_Of_Total = share of SDG claims in total text; (5) Sent_Opp_Of_EnvContent = opportunity sentiment within environmental content; (6) Sent_Risk_Of_EnvContent = risk-related sentiment within environmental content; (7) SDG_Sent_Opp_Of_SDGContent = opportunity sentiment within SDG content; (8) SDG_Sent_Risk_Of_SDGContent = risk-related sentiment within SDG content.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1)	1.00	0.71	0.50	0.48	-0.01	-0.13	0.10	0.04
(2)	0.71	1.00	0.43	0.51	0.37	-0.32	0.12	-0.08
(3)	0.50	0.43	1.00	0.87	0.18	-0.09	0.17	-0.04
(4)	0.48	0.51	0.87	1.00	0.16	0.08	0.20	-0.06
(5)	-0.01	0.37	0.18	0.16	1.00	-0.45	0.18	-0.12
(6)	-0.13	-0.32	-0.09	0.08	-0.45	1.00	-0.01	0.25
(7)	0.10	0.12	0.17	0.20	0.18	-0.01	1.00	-0.16
(8)	0.04	-0.08	-0.04	-0.06	-0.12	0.25	-0.16	1.00

7.1 Multicollinearity check for the ESM model

Table 10: Pearson Correlation Matrix

Variable	Size	ESG Score	ESM
Size	1.00		
ESG Score	0.34***	1.00	
ESM	0.10**	0.31***	1.00

Notes: Significance levels: * $p < 0.05$, ** $p < 0.01$,
*** $p < 0.001$.

Table 11: Multicollinearity Diagnostics

Variable	VIF	Tolerance
Size	1.13	0.88
ESG Score	1.24	0.81
ESM	1.11	0.90

Note: Variance Inflation Factors computed from full logistic models (Tab 4 and Tab 5). All values remain substantially below the conservative threshold of 2.5 (Allison, 2012).

The Pearson correlation analysis reveals statistically significant relationships among the core explanatory variables. Size demonstrates a moderate positive correlation with ESG scores ($\rho = 0.34$, $p < 0.001$), while showing a weaker but statistically significant association with environmental sentiment metrics ($\rho = 0.10$, $p = 0.013$). The correlation between ESG performance and environmental sentiment ($\rho = 0.31$, $p < 0.001$) suggests conceptual alignment while maintaining empirical distinctness.

The variance inflation factors (VIF) presented in Table 11 confirm the absence of problematic multicollinearity in our full specifications. Diagnostic statistics were computed from actual logit models with dependent variables ESGCs and GWAs at time t and $t + 1$. All VIF values remain well below 1.3, substantially under the conservative threshold of 2.5 recommended for logistic regression (Allison, 2012). Tolerance statistics between 0.81-0.90 exceed the critical value of 0.4 Menard, 2001, indicating that each predictor contributes unique explanatory power.

7.2 Robustness Tests

Several robustness checks were conducted to validate the primary findings. Table 12 presents the firm-level fixed effects estimates for the SDGOI model, incorporating Firm, Year, Sector, Domicile, and Report fixed effects. The coefficient estimates remain substantially consistent with the baseline results reported in the main analysis.

To further assess robustness, a fractional logit model with correlated random effects (Papke & Wooldridge, 2008) was estimated, the results of which are shown in Table 13. This alternative specification confirms that the positive association with opportunity-oriented sentiment is robust to different modelling assumptions concerning the dependent variable’s distribution and the treatment of unobserved heterogeneity.

Additionally, to mitigate the potential influence of outliers in the text-derived disclosure metrics, all regressors were winsorized at the 1st and 99th percentiles. The regression results using these winsorized variables, corresponding to the specification in Table 1, are reported in Table 14. The findings are qualitatively unchanged, indicating that the core results are not driven by extreme values in the textual data.

As a placebo test, we randomize the opportunity-oriented environmental sentiment component used to construct the ESM and recompute the index accordingly. Results are reported in Table 15. Across nearly all specifications, the randomized ESM (i.e. Rand_ESM) is statistically insignificant with respect to ESG controversies (ESGCs) or greenwashing accusations (GWAs). In contrast, firm Size and ESG Scores remain positive and highly significant across all models, as in the main analysis.

Overall, these results confirm that the content of the original, non-randomized ESM does not arise mechanically from the structure of the index nor from generic sentiment variation. When the sentiment component is randomized, the signal disappears, reinforcing the conclusion that the text-derived ESM captures economically meaningful information about managerial environmental messaging.

Table 12: Firm-level fixed effects estimates. Standard errors clustered at firm level.

	(1) SDGOI	(2) SDGOI	(3) SDGOI	(4) SDGOI
EnvContent_Of_Total	0.007 (0.136)			
EnvClaims_Of_Total		0.157 (0.391)		
SDGContent_Of_Total			-4.893*** (1.040)	
SDGClaims_Of_Total				-6.629*** (2.129)
Sent_Opp_Of_EnvContent	0.289** (0.145)	0.275* (0.153)	0.329** (0.142)	0.330** (0.132)
Sent_Risk_Of_EnvContent	-0.170 (0.313)	-0.156 (0.314)	-0.146 (0.306)	-0.183 (0.300)
SDG_Sent_Opp_Of_SDGContent	-0.429*** (0.066)	-0.430*** (0.066)	-0.413*** (0.066)	-0.416*** (0.064)
SDG_Sent_Risk_Of_SDGContent	-0.129 (0.130)	-0.127 (0.129)	-0.153 (0.134)	-0.151 (0.130)
Bond_Activi	0.001** (0.000)	0.001** (0.000)	0.001* (0.001)	0.001 (0.001)
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Sector FE	Yes	Yes	Yes	Yes
Domicile FE	Yes	Yes	Yes	Yes
Report FE	Yes	Yes	Yes	Yes
Observations	667	667	667	667

Note: Standard errors clustered at firm level in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 13: Fixed effects correlated random effects (CRE) fractional logit estimates. Standard errors clustered at issuer level.

	(1) SDGOI	(2) SDGOI	(3) SDGOI	(4) SDGOI
Intercepts	3.564*** (0.827)	3.780 *** (0.857)	3.469*** (0.875)	2.597*** (0.865)
EnvContent_Of_Total	-0.003 (0.699)			
EnvClaims_Of_Total		0.762 (2.074)		
SDGContent_Of_Total			-35.711*** (9.415)	
SDGClaims_Of_Total				-57.044*** (19.115)
Sent_Opp_Of_EnvContent	1.700** (0.804)	1.646* (0.830)	1.604** (0.774)	2.127** (0.815)
Sent_Risk_Of_EnvContent	-1.655 (2.150)	-1.513 (2.181)	-1.928 (2.203)	-0.984 (2.244)
SDG_Sent_Opp_Of_SDGContent	-2.249*** (0.376)	-2.261*** (0.380)	-2.049*** (0.354)	-2.050*** (0.382)
SDG_Sent_Risk_Of_SDGContent	-0.423 (0.613)	-0.414 (0.612)	-0.483 (0.582)	-0.693 (0.565)
Bond_Activi	0.002 (0.0005)	0.002 (0.0005)	0.001 (0.0005)	0.000 (0.0005)
Firm FE (via Mundlak)	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Sector FE	Yes	Yes	Yes	Yes
Observations	667	667	667	667

Note: Standard errors clustered at firm level in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Results are also robust to alternative operationalization of opportunity-oriented environmental tone. When replacing ESM with the ratio-based metric Sent_Opp_Of_EnvContent, the coefficient follow the same pattern as ESM. Effect sizes are quantitatively comparable to those obtained with ESM, and model fit (pseudo- R^2 , log-likelihood, information criteria) changes only marginally. These findings confirm that the ESM signal is not driven by report length or the choice of a count-based measure. Full estimation outputs are available upon request.

Table 14: Robustness: Fixed-Effects Estimates with Winsorised Regressors (OLS). Findings remain robust when also including Domicile and Type of Reports as additional control variables.

	(1) SDGOI	(2) SDGOI	(3) SDGOI	(4) SDGOI
EnvContent_Of_Total	0.027 (0.122)			
EnvClaims_Of_Total		0.148 (0.386)		
SDGContent_Of_Total			-5.208*** (1.038)	
SDGClaims_Of_Total				-7.247*** (2.319)
Sent_Opp_Of_EnvContent	0.280** (0.141)	0.264* (0.147)	0.315** (0.130)	0.314** (0.138)
Sent_Risk_Of_EnvContent	-0.033 (0.029)	-0.323 (0.294)	-0.346 (0.288)	-0.316 (0.287)
SDG_Sent_Opp_Of_SDGContent	-0.427*** (0.065)	-0.428*** (0.065)	0.413*** (0.063)	-0.409*** (0.065)
SDG_Sent_Risk_Of_SDGContent	-0.305* (0.174)	-0.297* (0.175)	-0.318* (0.170)	-0.335* (0.175)
Bond_Activi	0.001* (0.000)	0.001* (0.000)	0.000 (0.000)	0.000 (0.000)
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Sector FE	Yes	Yes	Yes	Yes
Observations	667	667	667	667

Notes. This table reports firm-fixed-effects estimations of the SDG Omission Index (SDGOI) using regressors winsorised at the 1% level. Winsorisation is a standard robustness procedure to reduce the influence of extreme values in text-derived disclosure metrics, which are known to exhibit heavy tails and leverage effects (see, e.g., Hadi, 1992; Huber, 1981; Rousseeuw and Leroy, 2003). Applying winsorisation to the environmental and SDG-related content and sentiment measures ensures that findings are not driven by outliers arising from unusually long reports, atypical linguistic structures, or isolated spikes in disclosure intensity. We re-estimate the baseline fixed-effects model substituting all continuous regressors with their winsorised counterparts. Results remain consistent in magnitude, sign, and significance across specifications, confirming that the main associations are not the product of outlier sensitivity.

Table 15: Placebo test using randomized environmental sentiment. Logistic regressions of ESG controversies (ESGCs) and greenwashing accusations (GWAs) at time t . The opportunity-oriented sentiment component of the ESM (i.e. Rand_ESM) is randomized across environmental sentences, and the ESM formula is recomputed accordingly. The models include year and sector fixed effects and standard errors clustered at the issuer level.

Dependent Variable:		ESGC _{<i>t</i>}			GWA _{<i>t</i>}	
<i>Variables</i>						
Rand_ESM	0.144 (0.107)	0.129 (0.119)	0.145 (0.122)	0.205* (0.105)	0.182 (0.120)	0.181 (0.121)
ESG Score		0.064*** (0.012)	0.059*** (0.012)		0.056*** (0.014)	0.052*** (0.014)
Size		0.372*** (0.112)	0.372*** (0.111)		0.273** (0.117)	0.269** (0.119)
ESM	0.670*** (0.194)		0.358** (0.166)	0.392*** (0.130)		0.179 (0.132)
<i>Fixed-effects</i>						
Year	Yes	Yes	Yes	Yes	Yes	Yes
Sector	Yes	Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>						
Observations	623	623	623	502	502	502
Squared Correlation	0.155	0.280	0.292	0.111	0.132	0.136
Pseudo R ²	0.091	0.202	0.211	0.045	0.100	0.098

Clustered (IssuerID) standard errors in parentheses

*Signif. Codes: *** 0.01, ** 0.05, * 0.1*