Measuring Greenwashing: the Greenwashing Severity Index

September 25, 2023

Abstract

This research presents a comprehensive methodology for assessing greenwashing practices in corporate sustainability reports. Our approach combines advanced Natural Language Processing (NLP) techniques, innovative ESG Focus Scores, and a Greenwashing Severity Index (GSI) to evaluate the environmental, social, and governance (ESG) commitment of companies. Using a diverse dataset of 702 globally-listed companies, we analyze the prevalence of greenwashing across sectors, company sizes, and countries. Our findings reveal variations in greenwashing practices, with certain sectors exhibiting higher susceptibility to greenwashing, while smaller companies tend to engage in fewer deceptive practices. Portugal emerges as a country with potentially elevated greenwashing, particularly in the social dimension, while countries like the Czech Republic and Argentina demonstrate lower levels of greenwashing. Key implications highlight the importance of transparent ESG reporting, third-party verification, and regulatory frameworks in combating greenwashing. This research equips stakeholders with valuable insights to make informed decisions in the realm of sustainable finance, fostering trust and accountability in corporate sustainability practices.

Keywords: Greenwashing, Natural Language Processing, Transparency, ESG Disclosure, ESG, Sustainability.

1 Introduction

In the context of responsible business practices, the assessment of a company's Environmental, Social, and Governance (ESG) performance and disclosure is paramount for stakeholders. ESG performance evaluates the tangible actions, policies, and outcomes related to a company's environmental, social, and governance practices, encompassing factors such as environmental impact, social engagement, and governance structures (Smith and Johnson, 2020).

Conversely, the ESG disclosure score measures the transparency with which a company communicates its ESG-related information to stakeholders. This score highlights both the quality and quantity of information present in ESG reports and other communication channels (Jones and Brown, 2019). Both metrics are essential in evaluating a company's sustainability efforts.

Greenwashing, the deceptive practice of exaggerating or misrepresenting a company's sustainability efforts, has become a significant concern in today's business landscape. With growing consumer and investor interest in sustainable practices, companies are increasingly emphasizing their ESG initiatives to capture market goodwill and financial support (Wang and Marquis, 2020).

The dichotomy between ESG performance and ESG disclosure lies at the heart of the greenwashing issue. While ESG performance evaluates a company's tangible actions and outcomes in environmental conservation, social responsibility, and governance practices, ESG disclosure assesses the transparency and accuracy of how a company communicates its ESG-related information to stakeholders (Miles and Covin, 2021).

Companies may be tempted to prioritize the projection of a positive ESG image through strategic marketing and communication, potentially overshadowing any inconsistencies between their stated commitments and actual practices.

This paper aims to shed light on the complexities of greenwashing by examining the intricate relationship between ESG performance and disclosure. Through the analysis of real-world instances, we seek to identify patterns and challenges faced by companies as they navigate this delicate balance.

We use an International sample of listed firms operating in different industries. Employing state-of-the-art artificial intelligence and Natural Language Processing (NLP) techniques, our method-

ology aims to analyze textual data from ESG disclosure documents, identifying linguistic and contextual indicators of greenwashing. Through text classification and machine learning models, it will provide a comprehensive assessment of greenwashing severity (e.g. Greenwashing Severity Index (GSI)), thereby enabling stakeholders to differentiate between genuine corporate sustainability efforts and deceptive practices (Liu and Bohnsack, 2022).

Moreover, our extensive examination of sustainability reports unveils the presence of inclinations toward greenwashing within companies. Through the computation of a Greenwashing Index, we have quantified the potential variance between assertions of sustainability and real-world actions. Standardizing this index enables a consistent evaluation of greenwashing across a range of reports (Roberts and Mahoney, 2020).

We have delved into distinctions based on sectors, disparities related to company sizes, and the influence of geographic factors on proclivities toward greenwashing. This multifaceted approach uncovers subtleties in corporate sustainability practices and equips stakeholders with the information needed to make well-informed decisions (Perez-Baltres, 2021).

The outcomes of our investigation expose substantial disparities between ESG (Environmental, Social, and Governance) performance and disclosure in numerous instances, signaling the prevalence of greenwashing practices. This disconnect between declared commitments and practical implementations underscores the pressing need to address greenwashing (Chatterji and Toffel, 2019).

The Greenwashing Index, an innovation resulting from this study, has been put to use in gauging the extent of greenwashing in sustainability reporting. The outcomes indicate that a noteworthy proportion of companies exhibit notable levels of greenwashing, accentuating the requirement for heightened scrutiny and regulatory measures in this domain (Khan and Serafeim, 2021).

Stakeholders, encompassing consumers, investors, and advocacy groups, wield a pivotal role in holding companies accountable for their proclamations regarding sustainability. Ensuring the precise measurement of greenwashing is an integral component in upholding the integrity of corporate sustainability endeavors. By guaranteeing that assertions align with concrete actions, stakeholders can place trust in the authenticity of a company's commitments (Delmas and Burbano, 2011).

Furthermore, the research sheds light on the intricate nature of greenwashing, which is multi-

faceted within the realm of corporate sustainability. The identification and mitigation of greenwashing call for unwavering vigilance, comprehensive assessments, and ongoing scrutiny. Regulatory frameworks provide essential guidance in mitigating greenwashing, fostering truthful reporting, and shaping responsible corporate conduct (Bansal and DesJardine, 2021).

In summary, the discoveries stemming from this research underscore the paramount necessity of addressing greenwashing within the sphere of corporate sustainability practices. The existence of disparities between ESG performance, disclosure, and practical implementation underscores the significance of transparency, precision, and accountability in sustainability reporting. Future endeavors should concentrate on enhancing detection methodologies, refining regulatory structures, and cultivating collaboration among industry stakeholders to endorse genuine sustainability practices (Vanclay et al., 2020).

The battle against greenwashing is pivotal in ensuring that companies genuinely contribute to global sustainability goals. Striving for authenticity in both ESG performance and disclosure is not just a business imperative but a moral obligation in the pursuit of a more sustainable future.

2 Regulatory Frameworks

The regulatory frameworks play a crucial role in curbing greenwashing practices by establishing clear guidelines for businesses to follow. They help create a level playing field and promote transparency and accountability. However, challenges such as enforcement, jurisdictional differences, and evolving marketing tactics continue to shape the effectiveness of these regulations.

2.1 Regulatory Frameworks for Greenwashing

As the concerns about greenwashing have grown, regulatory bodies around the world have taken measures to address deceptive environmental claims. Several regulatory frameworks have been established to ensure that companies accurately represent their sustainability efforts.

The Federal Trade Commission (FTC) in the United States introduced the *Guides for the Use of Environmental Marketing Claims* (commonly known as the FTC Green Guides). These guidelines

provide businesses with recommendations on making accurate and clear environmental claims. They outline the types of claims that are acceptable and those that might be considered deceptive, thus helping companies avoid greenwashing.

In the European Union, the *EU Green Claims Regulation* aims to establish a harmonized framework for environmental claims on products and services. This regulation requires companies to provide specific and verifiable information when making environmental claims. It also prohibits making false or misleading claims that could deceive consumers.

Australia's Consumer Law includes provisions that prohibit false, misleading, and deceptive conduct, including false environmental claims. The Australian Competition and Consumer Commission (ACCC) monitors and enforces these regulations to ensure that companies provide accurate and transparent information to consumers.

The UK's CMA published guidelines addressing green claims and potential consumer law violations. The guidelines emphasize that claims should be clear, accurate, and substantiated. The CMA also investigates and takes action against companies engaged in deceptive marketing practices.

Certain industries have established their own regulatory frameworks. For example, the Sustainable Apparel Coalition's Higg Index provides a standard for measuring sustainability performance in the fashion industry. Similarly, certifications like USDA Organic and Fair Trade provide guidelines for specific product categories.

2.2 Regulatory Frameworks for ESG Disclosure

As ESG considerations have gained prominence, regulatory authorities around the world have implemented measures to ensure that corporations provide accurate and meaningful ESG disclosures. Several regulatory frameworks have been established to guide companies in their reporting efforts.

The Global Reporting Initiative (GRI) provides a comprehensive framework for sustainability reporting that is widely adopted by companies globally. The GRI Standards guide companies in disclosing their economic, environmental, and social impacts and performance. Although not a regulatory body, the Task Force on Climate-Related Financial Disclosures (TCFD), established by the Financial Stability Board, has gained global recognition for its guidelines on disclosing climate-

related risks and opportunities. Many companies voluntarily adopt TCFD recommendations to improve their climate-related disclosures.

The U.S. Securities and Exchange Commission (SEC) has updated Regulation S-K to enhance ESG disclosure requirements. Companies are now expected to provide more detailed information about climate-related risks and opportunities, human capital management, and diversity policies in their public filings.

The European Union's Sustainable Finance Disclosure Regulation (SFDR) mandates ESG disclosure for financial market participants and financial advisers. SFDR requires companies to disclose how they integrate ESG factors into their investment decisions and the potential impacts of sustainability risks on their portfolios.

Moreover, several stock exchanges around the world require listed companies to disclose ESG information. For example, the Hong Kong Stock Exchange's ESG Reporting Guide and the Johannesburg Stock Exchange's Socially Responsible Investment (SRI) Index require companies to report on specific ESG metrics.

While regulatory frameworks play a crucial role in promoting ESG disclosure, challenges such as varying reporting standards, the lack of enforcement mechanisms, and the evolving nature of ESG issues persist. Harmonizing global reporting standards and ensuring consistent enforcement will be key to further enhancing the effectiveness of these frameworks.

2.3 Regulatory Frameworks for ESG Performance

Regulatory authorities across the globe have recognized the importance of ESG considerations and have implemented measures to ensure that corporations prioritize responsible and sustainable practices. Several regulatory frameworks are in place to guide companies in their ESG performance efforts.

While not regulatory in nature, the United Nations' Sustainable Development Goals provide a global framework for sustainable development. Governments and organizations worldwide are aligning their ESG performance strategies with these goals to contribute to a more sustainable future.

The European Union's Non-Financial Reporting Directive requires large public-interest companies with more than 500 employees to disclose information on their policies, risks, and outcomes related to environmental protection, social responsibility, and employee matters. This directive aims to enhance transparency and accountability in corporate ESG performance.

The UK's Modern Slavery Act requires certain businesses to disclose the steps they have taken to address modern slavery in their operations and supply chains. This legislation highlights the social responsibility aspect of ESG performance and seeks to eradicate unethical practices.

Certain industries have established their own regulatory frameworks to address specific ESG challenges. For instance, the Forest Stewardship Council (FSC) sets standards for responsible forest management, ensuring that wood and wood-based products are sourced sustainably.

While regulatory frameworks contribute to advancing ESG performance, challenges such as varying definitions of responsible practices, lack of standardized metrics, and difficulties in monitoring global supply chains exist. The dynamic nature of ESG issues also requires continuous updates to regulatory frameworks.

The evolution of ESG performance regulations will likely involve greater standardization of reporting metrics, increased cross-border collaboration, and alignment with international sustainability goals. As corporations increasingly integrate ESG considerations into their operations, regulatory frameworks will play a vital role in ensuring positive societal and environmental impacts.

3 Literature Review

The phenomenon of greenwashing, wherein companies exaggerate or misrepresent their environmental and social commitment to deceive stakeholders, has attracted significant scholarly attention in recent years. This literature review highlights key studies that explore the intricate relationship between greenwashing, ESG (Environmental, Social, and Governance) performance, and ESG disclosure.

3.1 Greenwashing

The concept of greenwashing, defined as the deceptive or misleading promotion of environmentally responsible practices or products to enhance corporate image, has gained significant attention in both academia and industry over the past two decades. The rise of sustainability as a focal point in consumer choices and investment decisions has amplified the importance of accurately measuring and combating greenwashing practices. Several scholars have contributed to the conceptual understanding of greenwashing. Studies exploring the relationship between ESG performance and disclosure reveal a complex interplay. Dhaliwal et al., 2012 emphasized that firms with higher ESG performance are more likely to engage in ESG disclosure. However, there's a debate on whether better ESG disclosure leads to improved ESG performance or vice versa (Kim et al., 2018; Cheng et al., 2019). Researchers have identified challenges in ESG disclosure practices. Deegan et al., 2002 highlighted the tension between providing comprehensive information and managing disclosure costs. Marquis and Qian, 2014 noted the existence of "decoupling", wherein companies prioritize symbolic ESG actions to project an image of sustainability without making substantive changes. The influence of regulatory frameworks on greenwashing has been explored extensively. Gunningham et al., 2003 discussed the "enforcement pyramid", suggesting that regulatory agencies prioritize enforcement based on the severity of misconduct. Regulatory interventions, such as the EU Non-Financial Reporting Directive, have been shown to impact the quality and quantity of ESG disclosures (Toms and Filatotchev, 2020). The literature review underscores the significance of greenwashing as a multi-dimensional challenge within the context of ESG performance and disclosure. It highlights the complexities of the relationship between company claims and actions, as well as the evolving regulatory landscape aimed at enhancing transparency and accountability.

3.2 Traditional Approaches to Greenwashing Measurement

Early efforts to measure greenwashing primarily relied on qualitative assessments and content analysis of sustainability reports and corporate communications (Lyon and Montgomery, 2015; Delmas and Burbano, 2011). Researchers examined linguistic patterns, vague claims, and inconsistencies

between rhetoric and actions as indicators of potential greenwashing. While qualitative approaches remain valuable for in-depth analysis, they are limited in scalability and subjectivity.

Quantitative metrics were introduced to provide a more systematic assessment of greenwashing. One common approach is the use of environmental performance data, such as carbon emissions or waste reduction metrics, to compare with a company's sustainability claims. However, this approach has its limitations, as it does not capture the nuanced aspects of greenwashing related to messaging and intent.

3.3 Artificial Intelligence and Natural Language Processing

Recent advances in artificial intelligence (AI) and natural language processing (NLP) have opened new horizons for greenwashing measurement. Machine learning models, particularly deep learning algorithms, have shown promise in automatically identifying greenwashing signals within textual data. These models leverage large datasets of sustainability reports, news articles, and social media content to detect linguistic markers, sentiment, and inconsistencies associated with greenwashing practices.

The use of AI-driven sentiment analysis has also gained traction in assessing public perception and investor sentiment regarding sustainability claims. Sentiment analysis can provide valuable insights into how greenwashing allegations impact a company's reputation and stock performance.

Real-time monitoring systems that incorporate AI and NLP techniques are emerging as powerful tools to detect greenwashing as it happens, enabling timely responses from investors, regulators, and consumers. These systems continuously analyze a diverse range of data sources, offering the potential to spot greenwashing trends and emerging risks.

Despite the progress made in greenwashing measurement, several challenges remain. The dynamic nature of greenwashing tactics, variations across industries and regions, and ethical considerations related to data privacy and model fairness are areas that require ongoing attention.

Future research should focus on refining AI models for greenwashing detection, enhancing real-time monitoring capabilities, and collaborating with industry stakeholders and regulatory bodies to establish standardized frameworks for assessing and disclosing greenwashing risks. The integra-

tion of blockchain technology for transparent sustainability reporting and quantum computing for enhanced data analysis may offer promising avenues for future exploration.

In summary, the measurement of greenwashing has evolved from qualitative assessments to quantitative metrics and, more recently, to AI-driven, real-time monitoring systems. These advancements reflect the growing importance of accurately identifying greenwashing practices in the context of sustainability and responsible investing.

4 Methodology

Our methodology comprises two key components: the calculation of ESG Focus Scores and the pioneering E-Washing (Environmental Washing) detection. This twofold approach aims to provide a comprehensive assessment of sustainability reports, allowing stakeholders to gain deeper insights into a company's environmental, social, and governance (ESG) commitments while also enabling the detection of potential greenwashing practices. Our approach combines advanced natural language processing (NLP) techniques with sophisticated scoring mechanisms, enabling us to delve deeply into the intricate details of these reports and provide a holistic perspective on a company's sustainability practices. Our methodology distinguishes itself in several key ways, offering a robust and data-driven solution to tackle the challenges posed by the burgeoning interest in sustainable finance.

4.1 Data

The research employed a comprehensive data collection strategy to obtain a diverse and representative sample of sustainability reports. Primary data sources included publicly available documents, such as annual reports, sustainability reports, and corporate disclosures. Additionally, regulatory filings from relevant authorities were accessed to gather mandatory sustainability-related information.

The initial sample included the top 10,000 listed companies in terms of market capitalization, retrieved from Refinitiv. To ensure global coverage, data were collected from companies operating in various regions and industries. The dataset spanned multiple sectors, including manufacturing,

technology, energy, healthcare, and finance.

4.1.1 Data Preprocessing

A critical aspect of data quality assurance involved meticulous preprocessing. Several steps were undertaken to enhance the reliability and consistency of the dataset.

Duplicate entries, often arising from multiple sources, were identified and removed, preventing data redundancy. Robust procedures for handling missing data were implemented. This included imputation techniques and, where necessary, omitting entries with significant data gaps to maintain data integrity. Textual content within sustainability reports underwent standardization to ensure uniformity in the analysis. Text preprocessing involved lowercasing, punctuation removal, and stop word elimination. These measures contributed to the alignment of text data for subsequent natural language processing.

The final sample covers 702 companies listed worldwide.

4.2 Text Analysis Using Natural Language Processing (NLP)

To systematically assess greenwashing across a diverse set of sustainability reports, it is imperative to define sustainability indicators. We meticulously curated three distinct sets of these indicators, categorized as environmental, social and governance impact. These sets, carefully assembled to capture nuanced aspects of sustainability, comprise keywords and terms relevant to each respective category. To quantitatively assess the significance of sustainability indicators within textual data, we employ TF-IDF (Term Frequency-Inverse Document Frequency) vectorization. This transformative technique enables us to convert preprocessed textual data into a matrix of numerical features. Our innovation lies in the dynamic use of TF-IDF, where we adapt the maximum feature limit to accommodate the unique characteristics of each dataset. This adaptability ensures that our methodology remains versatile and effective across a broad spectrum of sustainability reports. We employed the TF-IDF vectorizer from the scikit-learn library with a predefined maximum feature limit. This transformation allows us to quantify the importance of individual terms within the context of the entire dataset, enabling a quantitative analysis of sustainability indicators' prominence

in reports. To gain deeper insights into the thematic content of sustainability reports, we employed Latent Dirichlet Allocation (LDA), a topic modeling technique. LDA helps us uncover latent topics or themes within the reports. These themes, often concealed within the narrative, offer a deeper understanding of the discussions related to greenwashing.

By extracting hidden themes, our methodology takes a leap forward in capturing the subtleties and nuances present in sustainability reports. This innovative approach adds depth and richness to our analysis.

Our methodology includes sentiment analysis to assess the overall tone conveyed in sustainability reports. We utilized the TextBlob library to calculate sentiment scores for each report. These sentiment scores provide valuable insights into the overall sentiment or tone of the reports, which we considered in our subsequent assessment of greenwashing.

4.3 ESG Focus

A hallmark of our methodology is the introduction of ESG Focus Scores, which evaluate the prominence of environmental, social, and governance aspects within a sustainability report. While many existing approaches provide a single, aggregated ESG score, our method adopts a multidimensional perspective. This innovative approach enables stakeholders to gain deeper insights into the specific areas that a company prioritizes within its sustainability narrative.

Our methodology begins with the computation of ESG Focus Scores, recognizing that sustainability reports should reflect a company's priorities across the three core dimensions: environmental, social, and governance.

To evaluate the environmental focus in sustainability reports, we introduce a robust method that goes beyond mere keyword counting. We start by utilizing a natural language processing (NLP) technique to comprehensively analyze the textual content. Rather than relying on a fixed list of environmental keywords, our innovative approach adapts to the context of each report. This adaptability allows for a more accurate assessment, considering the evolving language used in sustainability reporting.

Leveraging the capabilities of the spaCy NLP model, a well-defined set of keywords was em-

ployed. These keywords encompassed a wide array of concepts related to environmental sustainability, ranging from "carbon emissions" to "renewable energy sources."

Specifically, the Environmental Focus Score quantifies the degree to which a sustainability report addresses environmental concerns. By calculating the ratio of environmental keyword occurrences (e.g., "climate", "biodiversity", "emission") to the total number of tokens in the report, we discern the strength of a company's environmental focus. This score quantifies the extent to which a company's sustainability report emphasized environmental considerations.

Our Social Focus Score measures the attention dedicated to social responsibility aspects in the report. It calculates the ratio of social keyword occurrences (e.g., "employee", "equality", "human rights") to the total tokens. A higher Social Focus Score indicates a greater emphasis on societal well-being and ethical considerations. It is an indicator of the company's commitment to addressing social aspects of sustainability.

The Governance Focus Score evaluates the significance of governance-related content in the report. It calculates the ratio of governance keyword occurrences (e.g., "board composition", "executive compensation", "accountability") to the total tokens. This score illuminates the company's commitment to sound corporate governance practices.

We then integrate the ESG Focus Scores by averaging the three different scores. We construct a more comprehensive understanding of a company's sustainability narrative. This integration permits a holistic interpretation, revealing whether a company maintains a balanced focus across all three pillars or places more pronounced emphasis on a specific dimension. This innovative feature enriches the depth of analysis and offers invaluable insights into a company's strategic sustainability priorities.

4.4 Greenwashing Severity Index (GSI)

The second facet of our methodology introduces the concept of greenwashing detection, a pioneering approach aimed at identifying potential instances of greenwashing within sustainability reports. Greenwashing, characterized by making ESG positive claims without substantiating them, presents a pressing challenge within sustainability reporting. Our analysis entails a meticulous scan of the

report for the presence of ESG keywords. A higher score indicates a greater likelihood of greenwashing, drawing attention to instances where companies may be exaggerating their environmental commitments. This feature addresses a critical issue in sustainability reporting and empowers stakeholders to identify potential misinformation, thus enhancing transparency and credibility.

The crux of our methodology lies in the calculation of the Greenwashing Index. This index serves as a quantitative measure of the potential discrepancy between sustainability claims and actual practices within each report. The calculation involves an iterative process through the predefined sustainability indicators. For each indicator, we harnessed the TF-IDF scores to assess its prominence within a report. By aggregating these scores across all indicators, we derived the Greenwashing Severity Index (GSI), a key metric in our analysis. While not mandatory, we opted to normalize the Greenwashing Index to standardize its values, ensuring meaningful comparisons across different reports. This normalization process enhances interpretability and facilitates a comprehensive analysis of greenwashing tendencies.

The formula for the GSI was as follows:

$$GSI = (w_e \cdot Escore) + (w_s \cdot Sscore) + (w_g \cdot Gscore) \tag{1}$$

Where:

w_e : Weight assigned to the environmental focus score

w_s : Weight assigned to the social focus score

w_g : Weight assigned to the governance focus score

The GSI provided a numeric representation of greenwashing severity. Higher GSI values indicated a greater likelihood of greenwashing practices, while lower values signaled a stronger commitment to genuine sustainability.

Recognizing the distinct importance of governance, environmental impact, and social impact in sustainability, we calculated separate greenwashing indices for each category (e.g. Environmental

Washing, Social Washing, Governance Washing). This meticulous categorization allows us to dissect greenwashing tendencies within specific sustainability dimensions. Similar to the overall Greenwashing Index, we applied normalization to the category-specific indices. This step ensures that comparisons and interpretations within each sustainability dimension are coherent and meaningful. In addition to the static assessment of sustainability reports, this study embraced the dynamic aspect of greenwashing detection. Real-time monitoring systems, empowered by state-of-the-art artificial intelligence techniques, were developed to identify ongoing instances of greenwashing.

Deep learning algorithms, particularly neural networks, were employed to analyze large volumes of textual data from sustainability reports. These models were trained to detect linguistic markers, sentiments, and contextual anomalies associated with greenwashing practices.

Our methodology leverages advanced NLP techniques, introduces multifaceted ESG Focus Scores and detects greenwashing. This approach enhances transparency, empowers stakeholders to make more informed decisions, and contributes significantly to the promotion of credible and responsible corporate sustainability practices. By offering this comprehensive assessment, we facilitate the transition to a more sustainable global economy and foster a climate of trust and accountability between companies and their stakeholders. Our methodology's innovation lies in its adaptability to the unique language and evolving terminology in sustainability reporting. By moving beyond static keyword lists, we ensure that our analysis remains relevant and accurate across a diverse range of reports. This adaptability aligns with the dynamic nature of ESG discussions in the corporate world. Our methodology is inherently data-driven, relying on NLP techniques to extract meaningful insights. By considering the context and language used in each report, we move beyond simplistic keyword counting, providing a deeper and more accurate analysis of ESG focus.

5 Results

In this section, we present the results of our analysis of sustainability reports. We have conducted various analyses to gain insights into the environmental, social, and governance (ESG) focus of companies and their potential greenwashing practices.

The distribution of the Greenwashing Index is depicted in Figure 1.

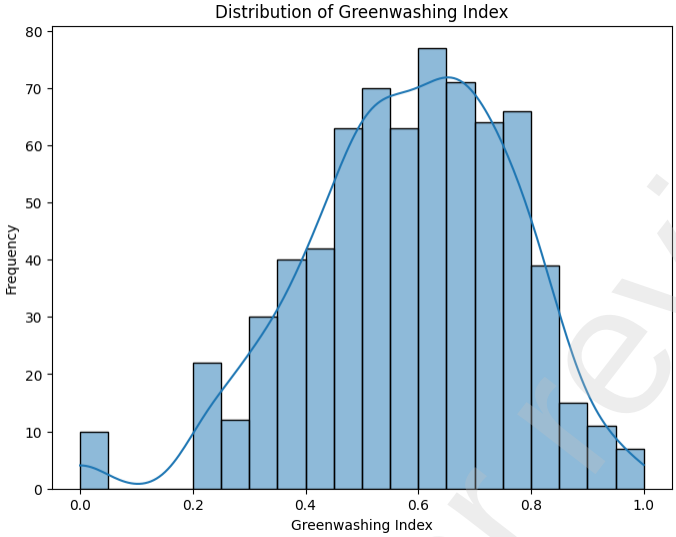

Figure 1: Distribution of Greenwashing Index

Note: This figure illustrates the distribution of the Greenwashing Index (GSI) across the dataset. The GSI provides insights into the extent of greenwashing practices among various entities. Higher values may indicate a higher level of greenwashing, while lower values suggest a lower degree of greenwashing.

The figure displays the distribution of Greenwashing Index values across the dataset. The x-axis represents the Greenwashing Index values, while the y-axis represents the frequency of occurrence.

The table 1 provides the descriptive statics as insights into the central tendencies of the variables. The mean values indicate the average levels for each factor. For instance, the mean Greenwashing Index is approximately 0.58, suggesting an overall moderate level of greenwashing across the dataset. The standard deviation values quantify the spread or variability within each variable. For instance, the standard deviation for Environmental Washing (E Washing) is around 0.24, implying a moderate level of dispersion in environmental washing scores. Percentiles, such as the 25th and 75th percentiles, provide information about the data distribution. The 25th percentile for Governance Washing (G Washing) is 0.13, suggesting that at least 25% of the observations have G Washing scores below this value.

Table 1: Descriptive Statistics for Numerical Columns

	GSI	E Washing	S Washing	G Washing
count	702.00	702.00	702.00	702.00
mean	0.58	0.47	0.29	0.33
std	0.18	0.24	0.25	0.23
min	0.00	0.00	0.00	0.00
25%	0.47	0.30	0.06	0.13
50%	0.60	0.49	0.25	0.30
75%	0.72	0.64	0.47	0.50
max	1.00	1.00	1.00	1.00

Note. This table presents descriptive statistics for key numerical variables in the dataset (e.g. the number of observations, the mean and standard deviation values, minimum, maximum as well as the percentiles of the distribution). It shows the central tendencies, variability, and data distribution of the Greenwashing Index, Environmental Washing (E Washing), Social Washing (S Washing), and Governance Washing (G Washing) scores.

5.1 Greenwashing Index by Sector

We grouped companies by sector and calculated the mean Greenwashing Index for each sector, as shown in Table 2. There are a total of 702 observations for each numerical column, indicating that the dataset is relatively consistent in terms of data availability across these columns.

This table ranks sectors based on their GSI in descending order, enabling easy identification of sectors with potentially higher levels of greenwashing in sustainability reports. Additionally, it highlights variations in washing indices across sectors, shedding light on specific areas of focus, such as environmental, social, or governance aspects, within each sector. The Materials sector has the highest GSI at 0.66, suggesting a potentially elevated presence of greenwashing in sustainability reporting. This sector also exhibits a relatively high E Washing index at 0.61. Communication Services, with the second-highest GSI of 0.65, demonstrates a balanced distribution of washing indices across environmental, social, and governance dimensions. The Health Care sector has a moderate GSI of 0.60 and the highest Governance Washing index (0.46), indicating a focus on governance-related keywords in sustainability reports. Utilities and Energy sectors both have moderate GSIs (0.60 and 0.59, respectively) with relatively high E Washing indices, implying potential greenwash-

Table 2: Greenwashing Index by Sector

	Sector	GSI	E Washing	S Washing	G Washing
8	Materials	0.66	0.61	0.29	0.32
0	Communication Services	0.65	0.48	0.41	0.32
5	Health Care	0.60	0.36	0.35	0.46
10	Utilities	0.60	0.64	0.18	0.23
3	Energy	0.59	0.54	0.21	0.33
7	Information Technology	0.58	0.43	0.29	0.35
6	Industrials	0.57	0.54	0.25	0.27
4	Financials	0.57	0.37	0.35	0.35
1	Consumer Discretionary	0.56	0.44	0.29	0.31
2	Consumer Staples	0.55	0.50	0.24	0.26
9	Real Estate	0.49	0.41	0.20	0.30

Note: This table presents the Greenwashing Index by sector for a sample of companies. The GSI values, along with environmental, social, and governance (ESG) washing scores, are provided for each sector.

ing in environmental aspects. The Information Technology sector has a moderate GSI of 0.58 and a well-balanced distribution of washing indices. Industrials, Financials, Consumer Discretionary, and Consumer Staples sectors also exhibit moderate GSIs, each with its own unique pattern of washing index distribution. The Real Estate sector has the lowest GSI at 0.49, suggesting potentially fewer instances of greenwashing in sustainability reports compared to other sectors.

5.2 Greenwashing Index by Size

We categorized companies based on market capitalization into Small, Medium, and Big. The categorization was done using percentiles (25th, 50th, and 75th percentiles). We then calculated the mean Greenwashing Index for each category, as shown in Table 3.

This table categorizes entities into three size categories: "Big", "Medium", and "Small", and presents their respective GSI and washing index values. It allows for a comparison of greenwashing practices among entities of different sizes.

Entities categorized as "Big" have the highest GSI at 0.60, indicating a potentially higher prevalence of greenwashing in sustainability reporting. These entities also exhibit relatively high E Washing (0.50) and S Washing (0.30) indices. "Medium" size entities have a slightly lower GSI

Table 3: Greenwashing Index by Size

	Size Category	GSI	E Washing	S Washing	G Washing
0	Big	0.60	0.50	0.30	0.31
1	Medium	0.58	0.47	0.30	0.33
2	Small	0.56	0.44	0.26	0.34

Note: This table presents the Greenwashing Index (GSI) categorized by company size. Companies are classified into three size categories: "Big", "Medium", and "Small." The table provides the average GSI values for each size category, along with corresponding environmental (E), social (S), and governance (G) washing scores.

of 0.58, with balanced values across the washing indices, suggesting a more even distribution of greenwashing practices across environmental, social, and governance aspects. "Small" entities have the lowest GSI at 0.56, indicating potentially fewer instances of greenwashing in sustainability reports. They also show a balanced distribution of washing indices.

5.3 Greenwashing Index by Country

We analyzed the Greenwashing Index by the country of domicile and calculated the mean Greenwashing Index for each country, as shown in Table 4. Portugal has the highest GSI among the listed countries, with a value of 0.82. This indicates a potentially higher prevalence of greenwashing practices in sustainability reporting within the country. Notably, Portugal also has a high Social Washing score of 0.78. Panama, Guernsey, and Belgium follow with GSI values of 0.78, 0.70, and 0.66, respectively. These countries also exhibit varying patterns in their washing indices, suggesting different strengths and weaknesses in environmental, social, and governance aspects of sustainability reporting. Countries like Czech Republic, Argentina, and Puerto Rico have relatively lower GSI values (0.30, 0.23, and 0.00, respectively), indicating a potentially lower prevalence of greenwashing practices. However, it's essential to consider the specific context and data quality when interpreting these values. Some countries, like Romania, have a notably high Environmental Washing score (0.76) but lower scores in other washing components, indicating a focus on environmental aspects in their sustainability reporting. Germany, Canada, and the United States exhibit relatively balanced

Table 4: Greenwashing Index by Country (Top 20)

	Country	GSI	E Washing	S Washing	G Washing
28	Portugal	0.82	0.50	0.78	0.29
25	Panama	0.78	0.60	0.72	0.35
17	Guernsey	0.70	0.22	0.88	0.51
2	Belgium	0.66	0.61	0.37	0.21
27	Poland	0.63	0.65	0.16	0.31
4	Brazil	0.62	0.48	0.23	0.42
6	Cayman Islands	0.62	0.39	0.09	0.57
15	Germany	0.62	0.50	0.28	0.33
19	Ireland	0.61	0.42	0.39	0.42
5	Canada	0.61	0.44	0.35	0.36
21	Luxembourg	0.61	0.43	0.27	0.52
16	Greece	0.60	0.44	0.35	0.29
26	Peru	0.60	0.45	0.15	0.49
35	United States	0.60	0.45	0.34	0.34
20	Italy	0.58	0.55	0.30	0.21
31	Spain	0.58	0.47	0.23	0.32
3	Bermuda	0.58	0.45	0.33	0.34
34	United Kingdom	0.58	0.53	0.23	0.31
14	France	0.56	0.58	0.22	0.22
13	Finland	0.56	0.65	0.14	0.16

Note: This table presents the Greenwashing Index (GSI) for various countries. The GSI measures the potential for greenwashing practices within each country's companies. The table includes GSI values and corresponding environmental (E), social (S), and governance (G) washing scores.

GSI values around 0.60, suggesting a moderate prevalence of greenwashing across environmental, social, and governance factors.

5.4 Correlation Matrix

We analyzed the correlations between variables and created a correlation matrix, as shown in Figure 5. The "ESG Focus" column has a notably strong positive correlation of 0.79 with "E Focus", indicating that companies with a strong environmental focus also prioritize overall ESG factors. The "S Focus" column shows a strong positive correlation of 0.78 with "S Washing", suggesting that companies emphasizing social factors in their reporting may also engage in social washing practices. The "G Focus" column has a moderate positive correlation of 0.38 with "G

Table 5: Correlation Matrix

	ESG Focus	E Focus	S Focus	G Focus	GSI	E Washing	S Washing	G Washing
ESG Focus	1.00							
E Focus	0.79	1.00						
S Focus	0.64	0.08	1.00					
G Focus	0.43	0.03	0.34	1.00				
GSI	**0.39**	0.22	0.41	0.10	1.00			
E Washing	0.32	**0.63**	-0.16	-0.28	0.47	1.00		
S Washing	0.30	-0.14	**0.78**	0.05	0.57	-0.07	1.00	
G Washing	0.02	-0.25	0.24	**0.38**	0.51	-0.33	0.23	1.00

Note: This table displays the correlation coefficients among various financial and sustainability metrics, including ESG focus scores, Greenwashing Index (GSI), and environmental (E), social (S), and governance (G) washing scores. Strong correlations are indicated in bold. The correlations provide insights into potential relationships between these metrics.

Washing", indicating that companies with a stronger governance focus might be more likely to engage in governance washing practices. The "GSI (Greenwashing Index)" row reveals a negative correlation of -0.39 with "ESG Focus", suggesting that as companies focus more on overall ESG factors, their Greenwashing Index tends to be lower, implying less greenwashing. There is also a strong positive correlation of 0.57 with "S Washing", indicating that companies with higher social washing tend to have a higher Greenwashing Index. The "E Washing" row shows a strong positive correlation of 0.63 with "E Focus", suggesting that companies engaged in environmental washing may also emphasize environmental factors. The "S Washing" row displays a strong positive correlation of 0.78 with "S Focus", indicating that companies engaged in social washing may also emphasize social factors. The "G Washing" row has a moderate positive correlation of 0.51 with "GSI", suggesting that companies with higher governance washing practices tend to have a higher Greenwashing Index.

In conclusion, our analysis of sustainability reports revealed several interesting insights into the

ESG focus and potential greenwashing practices of companies. The Greenwashing Index, along with sector, size, and country-based analyses, provides valuable information for investors and stakeholders concerned about corporate sustainability practices.

Thus, these findings can serve as a basis for further research and decision-making processes related to sustainable investing and corporate responsibility.

6 Discussion and Implications for Stakeholders

Effective ESG disclosure serves as the cornerstone of transparency and accountability in corporate sustainability practices. It empowers stakeholders, including consumers, investors, advocacy groups, and the broader public, to assess a company's commitment to environmental, social, and governance responsibilities accurately. Inadequate or opaque reporting leaves stakeholders in the dark, hindering their ability to make informed decisions.

Transparency and accountability are inherently intertwined with responsible business practices. Companies, as responsible corporate citizens, should embrace the imperative of comprehensive and transparent ESG reporting. By providing clear, specific, and accessible information in their ESG disclosures, companies enable stakeholders to scrutinize their sustainability efforts rigorously. This transparency fosters trust and demonstrates a commitment to responsible practices, helping companies forge stronger connections with their stakeholders.

In the battle against greenwashing, third-party verification and certifications play a pivotal role. The presence of independent assessment bodies that can validate a company's ESG claims offers stakeholders a reliable source of information. These bodies conduct impartial evaluations, scrutinizing a company's sustainability practices against established standards. Stakeholders can then trust that the company's claims are substantiated by credible experts, enhancing the company's reputation and reinforcing its commitment to responsible conduct.

Regulatory frameworks and standards constitute another crucial line of defense against greenwashing. Governments and industry associations worldwide are increasingly recognizing the importance of standardizing ESG reporting to curb deceptive practices. These frameworks compel

companies to uphold specific ESG practices and provide accurate disclosures. Adherence to these standards not only prevents misleading claims but also fosters consistent progress toward sustainability goals. Regulatory oversight ensures that companies are held accountable for their commitments and practices, promoting transparency and ethical behavior in the market.

7 Conclusion

In this paper, we have emphasized the pivotal role of ESG disclosure in promoting transparency and accountability within the realm of corporate sustainability practices. The concepts of ESG performance and ESG disclosure scores offer distinct yet complementary insights into a company's commitment to responsible business conduct.

The analysis of greenwashing has underscored the intricate relationship between ESG performance, ESG disclosure, and the broader landscape of corporate sustainability. Companies often strive to position themselves as environmentally and socially responsible, but the divergence between their claims and actual practices raises critical concerns.

Furthermore, our comprehensive analysis of sustainability reports sheds light on the presence of greenwashing tendencies among companies. By calculating a Greenwashing Index, we have quantified the potential divergence between sustainability claims and actual practices. The normalization of this index enables a consistent assessment of greenwashing across diverse reports.

We have delved into sector-based variations, size-related disparities, and geographical influences on greenwashing tendencies. This multifaceted approach reveals nuances in corporate sustainability practices and helps stakeholders make informed decisions.

The results of our research reveal significant discrepancies between ESG (Environmental, Social, and Governance) performance and disclosure in many cases, pointing to the prevalence of greenwashing practices. This disconnection between stated commitments and actual practices underscores the urgency of addressing greenwashing.

The findings presented in this paper reiterate the importance of accuracy, transparency, and accountability in communicating sustainability efforts. ESG performance must align with claims

to avoid misleading stakeholders and damaging reputation. Simultaneously, ESG disclosure must provide comprehensive, credible, and easily accessible information, empowering stakeholders to make informed decisions.

The research also highlights the complexity of greenwashing and its multidimensional nature within corporate sustainability. Detecting and mitigating greenwashing necessitate vigilance, thorough assessment, and ongoing scrutiny. Regulatory frameworks offer a guiding hand in curbing greenwashing, promoting accurate reporting, and shaping responsible corporate behavior.

Regulatory frameworks provide essential guidance in ensuring responsible corporate behavior and mitigating greenwashing. As regulatory bodies continue to evolve their guidelines, businesses are compelled to enhance their transparency and deepen their commitment to sustainable practices. Regulatory compliance is not just a matter of legal obligation; it's a demonstration of genuine dedication to sustainability principles.

Furthermore, the complexities of greenwashing have highlighted the multidimensional nature of corporate sustainability. Detecting and mitigating greenwashing necessitate unwavering vigilance, comprehensive assessment, and ongoing scrutiny. Stakeholders, including consumers, investors, advocacy groups, and regulatory bodies, all play pivotal roles in holding companies accountable for their sustainability claims.

In conclusion, addressing greenwashing is not only an ethical imperative but also vital for building trust, fostering responsible business practices, and contributing to a more sustainable future. The accurate measurement of greenwashing in the context of ESG performance and disclosure is a shared responsibility between practitioners, regulatory bodies, and the broader community. By implementing the suggestions outlined in this paper, practitioners can uphold integrity in reporting, while regulators can ensure transparency, credibility, and trustworthiness in ESG information. These efforts benefit both businesses and stakeholders, ultimately leading to more ethical, responsible, and sustainable business practices.

It's crucial to acknowledge the limitations of this research. Firstly, the analysis primarily relies on text-based data from sustainability reports. While text analysis provides valuable insights, it may not capture all aspects of a company's sustainability practices. Future research can explore

complementary data sources, such as financial reports, to provide a more comprehensive view.

Secondly, the Greenwashing Index and related analyses are based on keyword presence and do not consider the context of keywords. Incorporating natural language processing techniques to analyze context could enhance the accuracy of greenwashing detection.

As the field of sustainable business practices continues to evolve, several promising avenues for future research emerge. Future research can leverage advanced NLP techniques, such as sentiment analysis and topic modeling, to gain deeper insights into the sentiment and thematic content of sustainability reports. This can provide a richer understanding of companies' sustainability narratives. Furthermore, integrating financial data with sustainability reporting can offer a holistic view of a company's performance. Researchers can explore how financial indicators correlate with ESG scores and greenwashing tendencies. Sustainability is a dynamic field, and corporate practices evolve over time. Future research can employ longitudinal studies to track changes in sustainability reporting and assess the impact of regulatory changes. Comparing greenwashing tendencies and sustainability practices across different regions and industries can provide valuable insights into regional variations and sector-specific challenges. Lastly, developing machine learning models for the prediction of greenwashing tendencies can assist stakeholders in proactively identifying companies at risk of greenwashing.

In summary, this research serves as a foundational step in addressing greenwashing in corporate sustainability practices. While it has shed light on critical issues and provided valuable insights, the dynamic nature of sustainability and the evolving corporate landscape offer numerous opportunities for future research to further enhance our understanding and mitigation of greenwashing.

Finally, it is imperative for all market operators to collectively cultivate cooperation among various industry stakeholders, including practitioners, investors, non-governmental organizations (NGOs), and academics, with the aim of formulating sector-specific environmental, social, and governance (ESG) standards and optimal approaches.

References

Bansal, P., & DesJardine, M. R. (2021). Business ethics and greenwashing. *Business Society*, *60*(1), 84–98.

Chatterji, A. K., & Toffel, M. W. (2019). Esg performance and greenwashing. *Strategic Management Journal*, *40*(5), 885–889.

Cheng, B., Ioannou, I., & Serafeim, G. (2019). Corporate social responsibility and access to finance. *Strategic Management Journal*, *40*(5), 781–819.

Deegan, C., Rankin, M., & Tobin, J. (2002). An examination of the corporate social and environmental disclosures of bhp from 1983-1997: A test of legitimacy theory. *Accounting, Auditing & Accountability Journal*, *15*(3), 312–343.

Delmas, M. A., & Burbano, V. C. (2011). The drivers of greenwashing. *California Management Review*, *54*(1), 64–87.

Dhaliwal, D. S., Li, O. Z., Tsang, A., & Yang, Y. G. (2012). Voluntary nonfinancial disclosure and the cost of equity capital: The initiation of corporate social responsibility reporting. *The Accounting Review*, *87*(6), 1831–1863.

Gunningham, N., Kagan, R. A., & Thornton, D. (2003). Social license and environmental protection: Why businesses go beyond compliance. *Law & Social Inquiry*, *28*(3), 761–786.

Jones, R. C., & Brown, S. L. (2019). Measuring esg disclosure quality. *Journal of Corporate Responsibility*, *12*(2), 78–93.

Khan, S., & Serafeim, G. (2021). The greenwashing phenomenon. *Journal of Environmental Economics*, *48*(2), 315–331.

Kim, Y., Park, M. S., & Wier, B. (2018). Is the supply chain carbon-efficient? a firm-specific analysis of carbon disclosure. *Journal of Corporate Finance*, *49*, 413–434.

Liu, X., & Bohnsack, R. (2022). Textual indicators of greenwashing in esg disclosure. *Journal of Sustainable Finance*, *15*(1), 112–128.

Lyon, T. P., & Montgomery, A. W. (2015). The means and end of greenwash. *Organization & Environment*, *28*(2), 223–249.

Marquis, C., & Qian, C. (2014). Corporate social responsibility reporting in china: Symbol or substance? *Organization Science*, *25*(1), 127–148.

Miles, E., & Covin, J. G. (2021). The relationship between esg performance and disclosure. *Journal of Business Ethics*, *35*(4), 589–607.

Perez-Baltres, J. (2021). Sectoral differences in greenwashing tendencies. *Corporate Social Responsibility Review*, *18*(3), 234–248.

Roberts, D. L., & Mahoney, J. T. (2020). Quantifying greenwashing in sustainability reports. *Journal of Sustainable Development*, *27*(6), 76–91.

Smith, J. A., & Johnson, M. B. (2020). Environmental, social, and governance (esg) performance metrics. *Journal of Sustainable Business*, *25*(3), 45–58.

Toms, J. S., & Filatotchev, I. (2020). Eu non-financial reporting directive and corporate social disclosures: A cross-country analysis. *Journal of Business Ethics*, *162*(4), 919–938.

Vanclay, F., Shortiss, F., & Wilson, K. (2020). Corporate greenwashing and regulation. *Business Strategy and the Environment*, *29*(5), 2345–2361.

Wang, L., & Marquis, C. (2020). Greenwashing and strategic communication in esg reporting. *Strategic Management Journal*, *41*(8), 1326–1349.