Louvain School of Management

How can artificial intelligence-driven language analysis models enhance the detection of greenwashing tactics in corporate ESG disclosures and what influence does this detection have on the stock market?

Author(s): Léa Dufrane
Supervisor(s): James Thewissen
Academic year 2024-2025
Dissertation for the master of Management (M120)
Master subject and focus on Corporate Finance
Daytime schedule

Abstract :

In recent years, Environmental, Social and Governance disclosures have become increasingly important for companies and investors worldwide. However, the credibility of these reports is often undermined by a phenomenon called "greenwashing". This occurs when firms exaggerate or misrepresent their sustainability efforts. This thesis explores how artificial intelligence, particularly language analysis models, can enhance the detection of greenwashing in corporate ESG disclosures, and how such detection influences investor behaviour.

A new Greenwashing Score is developed based on five key linguistic indicators: buzzword density, lack of specificity, sentiment tone, hedging language, and candor. This composite score is computed through rule-based natural language processing (NLP) techniques in R and challenged through an AI co-analyst using GPT-4. This gives a hybrid methodology that combines interpretability with advanced language understanding.

The second part of the study examines the financial impact of greenwashing by analyzing cumulative abnormal returns (CARs) following ESG report disclosures. Event study regressions test the relationship between greenwashing risk and stock market reactions across short- and medium-term event windows (CAR2, CAR5, and CAR60). Results show that higher greenwashing scores are to some extent significantly associated with negative investor responses, particularly when ESG credibility is questioned.

By linking AI-driven textual analysis to market outcomes, this research contributes to the growing field of sustainable finance and highlights the potential of AI tools in promoting transparency and accountability in ESG communication. It also provides investors and regulators with a data-driven method to identify potentially misleading disclosures and encourage more authentic sustainability reporting.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
Louvain School of Management
Place des Doyens, 1 bte L2.01.01, 1348 Louvain-la-Neuve
Boulevard Emile Devreux 6, 6000 Charleroi, Belgique
Chaussée de Binche 151, 7000 Mons, Belgique
www.uclouvain.be/lsm

Foreword

This thesis marks the end of a two-year journey within the Master's in Management program, and more broadly, of my academic path. It reflects both a personal interest in sustainable finance and the technical challenges of using AI to tackle real-world issues like greenwashing.

I would like to thank my thesis supervisor, Professor James Thewissen, for helping me identify and refine the topic at the early stage of this work. His guidance was instrumental in shaping the initial direction of this research.

Finally, I am grateful to my family and friends for their support, encouragement, and patience throughout this process, and to Sunset, my dog, for the calming company during long hours of writing.

Declaration

During the preparation of this master's thesis, the author utilized AI-based tools, including OpenAI's ChatGPT (GPT-4) and DeepL, for the following purpose:

1. *ChatGPT (GPT-4)* was used as a support tool in three areas:

• As a co-analyst, specifically in the context of the thesis's comparative design. The model was used to generate its own language-based Greenwashing Score (GWS_gpt), derived from full ESG disclosures, using its internal large language model capabilities. This AI-generated score was then compared to the author's independently constructed, rule-based score as part of the empirical analysis.

• To help review, clarify, and improve the structure and readability of the R code written by the author.

• As a writing assistant, especially for rephrasing, shortening overly long sections, correcting grammar, and improving clarity.

2. *DeepL (Write and Translator)* was used to enhance the fluency and readability of certain English passages.

3. After using OpenAI's ChatGPT (GPT-4) and DeepL, the author diligently reviewed and edited the content produced by the tool. We take full responsibility for the final content presented in this thesis.

By signing this declaration, we affirm that the content of this master's thesis reflects our original work, augmented by the responsible use of AI.

Date: 05/28/2025

Table of content

1. Introduction

In April 2025, one of Europe's largest asset managers learned the price of a "green" lie. Deutsche Bank's DWS was fined €25 million for exaggerating the sustainability of its investments. This reduced its stock value by around 2% in a single day (*Reuters*, 2025). This real-world example highlights a growing paradox in sustainable finance: while environmental, social and governance (ESG) principles are becoming more widely adopted, trust in corporate ESG claims is being affected by greenwashing scandals. Companies proudly promote their environmental credibility in eye-catching reports and press releases, yet stakeholders are increasingly questioning whether these claims reflect genuine action or are merely marketing tactics.

As one recent study pointed out, there is a tendency for more words to be used with less meaning. This serves as a reminder that an increase in ESG talk does not necessarily lead to real change (Hu et al., 2024). The stakes could not be higher. With over $30 trillion invested in assets labelled 'sustainable' in 2023, misleading ESG communication represents more than just a PR problem; it creates societal and financial risks (*Global Sustainable Investment Review Finds US$30 Trillion Invested in Sustainable Assets*, 2023). If sustainability reports become disconnected from reality and only serve as marketing tools, capital may be wasted and genuine progress may be held back. Regulators worldwide have recognised this danger and begun to take action against greenwashing, which is now treated as a form of fraud. Investors, too, are no longer indifferent.

If greenwashing is exposed, the market's response can be quick and severe. Akyildirim et al. (2023) found that firms accused of greenwashing saw average short-term stock drops of 0.6%. The authors concluded that breaches of trust in ESG reporting are interpreted as signals of poor management or governance. This leads to a negative evaluation by investors and a higher perceived risk. Similarly, Xu et al. (2025) observed consistent abnormal negative returns in response to greenwashing news, particularly among firms that had previously established themselves as ESG leaders. These findings confirm that a lack of credibility in ESG disclosure can have real financial consequences.

Within this context, artificial intelligence (AI) has emerged as a promising tool for improving transparency and rigour in ESG reporting. Research demonstrates that AI's ability to analyse complex datasets can improve the quality of disclosures and the accuracy of ESG ratings. This highlights its potential in the detection of greenwashing (T. Li et al., 2024; Zhang, 2024). Techniques such as natural language processing (NLP) enable AI-driven models to provide a reliable and objective method for identifying greenwashing tactics in corporate disclosures. By processing large volumes of text, NLP allows researchers to identify insights relating to sentiment, keyword usage and thematic emphasis that could suggest potentially misleading claims, which may not be apparent to human analysis. Tools such as GreenWatch.ai use NLP to compare the environmental claims made by companies with data from a variety of sources. They then identify inconsistencies and improve the accuracy of investor evaluations (*GreenWatch*, 2021; *Bloomberg*, 2021). AI has the potential to be a useful tool for improving the transparency and reliability of sustainability reporting (Tang et al., 2023).

This thesis looks at whether AI-powered language analysis models can spot greenwashing in how companies talk about their ESG efforts. In particular, it addresses the following research question: *How can artificial intelligence-driven language analysis models enhance the detection of greenwashing tactics in corporate ESG disclosures, and what influence does this detection have on the stock market?* To achieve this, we combine rule-based NLP techniques with financial event study methodology. We identify linguistic patterns such as excessive optimism, vague language, overused ESG buzzwords, and lack of candor. These are aggregated into a composite greenwashing score and benchmarked by industry. We then test whether these scores predict abnormal returns following disclosure. This provides insight into the market's sensitivity to ESG credibility.

A distinctive feature of this study is the integration of ChatGPT, a world-leading large language model (LLM), as a "virtual analyst". ChatGPT complements traditional, lexicon-based natural language processing by identifying more nuanced dimensions of narrative style and tone. This hybrid approach reflects the evolving frontier of AI applications in sustainability and finance. It represents an evolution in ESG analysis, moving beyond simple word counts to interpret how firms frame their sustainability efforts.

Beyond its methodological contribution, this thesis addresses a key financial issue: how do capital markets respond to disclosures that appear misleading or overly exaggerated? By linking AI-detected greenwashing risk scores to cumulative abnormal returns (CARs) over various time windows, we assess whether authenticity is rewarded and exaggeration penalised in both short- and medium-term market reactions.

Ultimately, this research aims to bridge the gap between what companies say and what investors believe. In a time when ESG reporting is under more scrutiny than ever before, the ability to use algorithms to check the authenticity of disclosures is valuable. If firms know their words can be analysed on a large scale, they may be less tempted to exaggerate or engage in greenwashing. On the other hand, if they are honest and balanced, they may gain the confidence of investors.

The thesis is organised as follows. The first section provides a review of the existing literature on greenwashing, the use of AI in corporate disclosures and market responses to ESG information, to give an overview of the subject. The research question and hypotheses are then presented. Following, it is the methodology section which details the construction of greenwashing scores using R and Chat-GPT, and describes the event study. The results section then presents statistical analyses testing the relationship between greenwashing indicators and CARs. Finally, we discuss findings and limitations encountered.

In an era characterised by an increasing focus on sustainability, the accuracy and reliability of ESG reporting is a key issue. This research contributes to the fields of finance, business ethics and sustainable investment by demonstrating the potential of AI to improve greenwashing detection and analysing its market impact. The findings could ultimately help stakeholders to better understand the challenges associated with ESG disclosures.

2. Literature Review

2.1. Evolution of ESG Disclosures and the Rise of Greenwashing

In recent decades, there has been a huge increase in the integration of environmental, social and governance considerations into corporate strategy and reporting. ESG reporting, once a niche practice, is now almost universal among large companies. For example, the proportion of S&P 500 firms issuing dedicated sustainability reports increased from around one-third in 2010 to 86% by 2021 (Rouen et al., 2022). This growth reflects the increasing importance of ESG information for investors, regulators and stakeholders. The value of assets under management in ESG-focused funds has grown to the point where it is now measured in trillions of dollars, and stakeholders are increasingly demanding transparency regarding the social and environmental impacts of corporations (Cifrino, 2023).

However, rising expectations also bring challenges. Given that firms have come to recognise that their image, legitimacy and reputation are at risk if they do not comply, they may be tempted to present their environmental actions in a more favourable light than is actually the case (Arouri et al., 2021). The objective is to persuade the public that their efforts are making a meaningful contribution to environmental sustainability, even though this may not be true (Kim et al., 2017). This unethical trend is what we call “Greenwashing”.

Greenwashing is essentially defined by scholars as a form of misrepresentation, whether by omission or exaggeration, whereby firms portray themselves as more sustainable than they truly are. A widely cited definition by Delmas and Burbano (2011) describes greenwashing as “the intersection of two firm behaviors: poor environmental performance and positive communication about environmental performance”. Lyon and Maxwell (2011) similarly define greenwashing as the selective disclosure of positive information about a company's ESG performance, without providing a full picture of the negatives. These definitions highlight that greenwashing is fundamentally a misalignment between the positive image a company presents and its actual, less sustainable practices.

Why do firms engage in greenwashing? Several external and internal motivations have been identified. Externally, companies face rising pressure from regulators, investors, consumers, and activists to demonstrate sustainability (Delmas & Burbano, 2011). In the absence of strict verification and enforcement, a firm may find it easier to claim good conduct than to actually improve it (Lyon & Maxwell, 2011). Greenwashing can therefore be an appealing short-term strategy designed to improve a company's reputation or avoid criticism without making the required investments in meaningful ESG improvements (Walker & Wan, 2012). This tendency is particularly common when regulations and monitoring are limited, which gives exaggerated claims little reason to be stopped (Toffel & Marquis, 2012).

Internal factors also play a role, such as organizational dynamics and incentives. For example, if a company’s sustainability performance is lacking, but its communication’s team excels at messaging, it may frame its ESG image positively (Christensen et al., 2013). At an individual level, managerial over-optimism and psychological mechanisms such as moral licensing can also contribute. This is the belief that previous good behaviour justifies current bad behaviour, and that can contribute to greenwashing (Merritt et al., 2010).

There are many ways in which greenwashing can be included in corporate communications, from the most subtle to the most obvious. Researchers distinguish firm-level greenwashing (portraying the entire organization as environmentally responsible) from product-level greenwashing (misrepresenting a specific product or service as "green") (de Freitas Netto et al., 2020) . At firm level, for example, a company may publish a sustainability report full of eco-friendly claims, while in reality, its core business continues to cause significant pollution. At the product level, a company may advertise a particular item with misleading environmentally friendly claims or labels. Previous studies and organisations such as consumer advocacy groups (e.g. Terrachoice, 2009) have identified a variety of common greenwashing tactics. Key tactics include:

- **Selective disclosure:** This involves highlighting positive environmental achievements while omitting negative information. For example, a mining company's report might emphasize land rehabilitation efforts or energy savings, while ignoring recent toxic spills and regulatory fines. By reporting only the positive aspects, the company creates a one-sided narrative of sustainability (Lyon & Maxwell, 2011).
- **Vague or misleading claims:** Ambiguous and meaningless buzzwords are used without any concrete evidence or definitions. Examples include describing a product as 'eco-friendly', 'sustainable' or 'all-natural' without any clear definition or certification. This kind of vague language plays on positive perceptions but does not offer any real facts (Delmas & Burbano, 2011). The overuse of feel-good buzzwords is a red flag that the communication may be more about marketing than actual facts.
- **Dubious labels or certifications:** Some companies claim to have third-party certifications that don't actually exist. Others have even invented their own green-sounding labels to impress stakeholders (de Freitas Netto et al., 2020). This tactic manipulates stakeholders' trust in certifications, when in reality no independent verification has actually been made.
- **Irrelevant or minor achievements:** Making a big deal of a small environmental improvement that does not address a company's overall impact (Terrachoice, 2009). For example, marketing a fundamentally polluting product with a 'biodegradable' wrapping. While the claim may be true, it does not relate to the product's main environmental issues.
- **Outright false claims:** In the worst cases, companies make false or unverifiable statements about their ESG performance. This could involve claiming reductions in emissions, clean technologies or eco-certifications that they do not actually have. Notorious examples include Volkswagen's "Dieselgate" scandal, in which the company marketed diesel cars as low-emission while secretly cheating emissions tests

(*VW "dieselgate" Fraud*, 2019). Legal action and public scandal can result from this kind of obvious misrepresentation (as seen in cases against Volkswagen, Keurig, and others).

All of these tactics have one thing in common: they all involve a discrepancy. This means there is a gap between symbolic communications (e.g. PR-driven claims, glossy reports, feel-good narratives) and substantive actions (e.g. actual ESG performance and outcomes). Essentially, greenwashing enables companies to benefit from the positive image associated with being 'green' without fully committing to the necessary work, as long as their claims remain unquestioned. As Lin et al. (2025) explain, firms may be tempted to make superficial ESG disclosures in order to maintain their reputation and competitive advantage. However, not all sustainability reports are examples of greenwashing; many firms use ESG disclosures to communicate their progress and challenges honestly. But some companies seem to use these reports rather to manage their image than to be accountable (Arouri et al., 2021). Although academic literature confirms the occurrence of greenwashing in ESG reporting, its extent is difficult to assess due to the subjective nature of perceived sincerity.

Such practices not only mislead stakeholders but also reduce the credibility of genuine sustainability initiatives, which results in a loss of public trust (Seele & Gatti, 2017). Moreover, greenwashing also hinders investors and analysts from obtaining reliable ESG data, complicating corporate accountability (Gehrig & Moreno, 2023). As a result, the growing awareness of these risks has forced stakeholders to become more vigilant. "Empty" green claims are more and more frequently spotted by consumers, with a negative response being shown by them once truth is revealed. For instance, marketing studies have found that when consumers discover that a company's environmental claims were misleading, their trust in the company and their intention to purchase its products drops significantly (De Jong et al., 2018). Greenwashing can therefore damage a brand's reputation and reduce customer loyalty, especially among the growing number of socially and environmentally conscious consumers (Amer & Ezz, 2023).

Regulators have also started to take action. In recent years, authorities in the US and EU have fined companies for making false sustainability claims, treating greenwashing as a form of consumer fraud (*European Commission*, n.d.; *Federal Trade Commission*, n.d.). These enforcement actions highlight that the consequences of such practices can include legal sanctions and public relations scandals.

2.2. Financial Market Reactions to Greenwashing and ESG Credibility

ESG disclosures serve as a communication channel to investors, and their credibility can impact investor behaviour and firm value. A central question for both academics and practitioners is: do capital markets reward sincere ESG efforts and punish greenwashing? While investors today pay close attention to corporate sustainability claims, they are also increasingly concerned about the quality and honesty of these claims.

Investors tend to react favourably when firms communicate meaningful, transparent ESG information. Moreover, the firm can benefit financially from good disclosure, which can reduce information asymmetry and perceived risk. For example, Dhaliwal et al. (2014) found that companies that published voluntary CSR reports had a consequently lower cost of equity capital, especially in jurisdictions with strong investor protection. This suggests that investors value the additional information, as they perceive firms that openly report sustainability data as less risky or more reliable. Similarly, Flammer (2013) showed that announcements of genuine environmental initiatives, such as emission reductions or certifications, trigger positive short-term stock price reactions.

Shareholders seem to reward firms that take real sustainability action, as they see this as an indicator of good management or potential future cost savings. Consistent and high-quality ESG reporting is associated with benefits such as higher firm valuation, improved financing efficiency, and enhanced stakeholder legitimacy (Habib et al., 2025). This can also lead to lower risk, as transparency reduces uncertainty (Chen et al., 2023). For instance, a recent study by Huang et al. (2024) reports that firms with more informative, data-driven sustainability disclosures attract a larger base of long-term investors and enjoy better stock liquidity. All these studies highlight a fundamental fact: markets value credibility. When a firm's ESG communication is perceived as accurate, specific and reliable, this can lead to tangible financial advantages such as a lower cost of capital or increased investor demand (Trinh et al., 2023).

On the other hand, suspected greenwashing can trigger negative market reactions. Once detected, it tends to result in financial losses. Indeed, when severe cases of greenwashing are exposed, they have been shown to trigger negative reactions in stock prices (Akyildirim et al., 2023). This market reaction reflects a loss of investor trust that may signal deeper management failures, leading shareholders to reassess the firm's value downward.

Further research validates these results. Xu et al. (2025) examine over 100 greenwashing news events in the post-2015 period (after the Paris Agreement). They also found a significant decline in firm value after the exposure of greenwashing. Furthermore, they noted that the market's reaction was more severe for companies that had previously claimed to have strong ESG practices. In other words, firms with strong prior reputations in this area are penalised more severely when they are exposed as greenwashers. This is probably because stakeholders feel betrayed or let down in such cases. Krüger (2015) found that investors penalise negative ESG events more severely than they reward positive news. This suggests that events that damage trust (such as a greenwashing scandal) have an outsized impact compared to announcements that build trust. A company that has promised more than it can deliver on sustainability may therefore see its stock price fall more quickly and for longer than the boost it got from the initial positive announcements.

It is important to recognize that both individuals and companies can face market penalties for greenwashing. And this punishment is not always linked to big lies that are discovered. Investors can also judge the sincerity of a company's disclosure in the absence of a scandal (Stuart et al., 2021). Companies that talk up sustainability but perform poorly tend to underperform financially, whether due to unmitigated risks becoming reality or reputational damage. A 2024 PwC Global Investor Survey found that 94% of investors believe that corporate sustainability reports contain unsupported claims (PricewaterhouseCoopers, 2024). This reflects a general scepticism towards such disclosures. This lack of trust can cause investors to doubt the credibility of the information presented, which could affect their investment decisions (Ernst et al., 2025).

That said, market reactions to greenwashing are not uniform and can vary depending on a range of contextual factors, including industry dynamics, regulatory environments, media influence, and firm-specific characteristics. Recent research highlights several important nuances in how these factors shape investor responses.

Industry context plays a significant role in determining the severity of market reactions. Investors tend to react more negatively to greenwashing in industries where environmental and social issues are particularly important. For example, one study found that sectors such as energy and mining, which have a significant environmental impact, experience a more significant decline in stock prices when greenwashing is exposed (Akyildirim et al., 2023).

These companies are under greater scrutiny, so stakeholders are less likely to forgive dishonesty.

The regulatory and cultural context also has a considerable influence. Firms in areas with strong environmental regulation and public awareness face steeper penalties. Akyildirim et al. (2023) find that companies caught greenwashing in these regions suffer larger stock price declines than those in less regulated environments. In places like the EU or California, where sustainability is prioritised, greenwashing is not only viewed as unethical, but also legally risky (Ballan & Czarnezki, 2024). Conversely, in markets where enforcement is weak or the concept of ESG is less established, investors may be more indifferent or slower to react.

Media exposure further amplifies market reactions. Studies indicate that when allegations of greenwashing receive considerable press coverage, particularly in reputable international channels, there is a tendency for negative stock returns to be more significant and persistent (Akyildirim et al., 2023). Media attention has a key role in amplifying the impact by increasing the number of informed investors and reinforcing concerns about a possible breach of trust. As a result, companies publicly portrayed as greenwashers often experience further declines in share price well beyond the initial market reaction.

A firm’s prior ESG track record also affects investor behaviour. Interestingly, a firm’s previous history of making ESG claims can influence investor reactions. If a company with an overall good reputation is exposed for greenwashing for the first time, investors tend to respond in shock, typically selling off shares quickly (Nicolas et al., 2023). By contrast, if a firm already has a history of controversial ESG behaviour, a new greenwashing incident may result in a less severe reaction. This is possibly because investors have already anticipated some level of dishonesty. In any case, the trend in recent years is that tolerance of greenwashing is decreasing. Since about 2010, it has become clear that markets have become stricter about greenwashing. This is because investors are more aware of ESG issues and want companies to be more transparent. (Akyildirim et al., 2023).

Expectations, in general, are a critical factor. Firms perceived as ESG leaders are held to higher standards by the market. Xu et al. (2025) provide evidence that, when a company with a high ESG rating is found to be greenwashing, the market reaction is more negative than it would be for a firm with a lower rating. This supports the idea that investors feel particularly ‘betrayed’ when supposedly sustainable firms' claims are found to be false. It also implies that they do consider external ESG ratings and information alongside the firm's own reports, and that discrepancies between the two can raise red flags. Indeed, Berg et al. (2022) highlight frequent divergences in ESG ratings (the so-called "aggregate confusion"), which can make it challenging to know which data to trust.

In sum, financial markets do not remain indifferent to greenwashing. Credible ESG disclosure can foster investor goodwill and potentially reduce the cost of capital. Conversely, once detected, greenwashing can result in real financial costs such as immediate stock price drops, reputational damage and higher risk premiums.

2.3. AI/NLP Techniques for Detecting Greenwashing in ESG Reports

Identifying greenwashing in corporate ESG disclosures is challenging, as it requires an assessment of the sincerity and completeness of textual content. Traditionally, the analysis of sustainability reports has been largely qualitative, with researchers either manually reading reports to spot inconsistencies, or using basic metrics such as word counts or sentiment ratios to infer credibility. Early content analyses revealed that many firms only report their achievements and omit their failures, which is viewed as a red flag for greenwashing. As Hahn and Lülfs (2014) point out, genuinely transparent reports tend to present a balanced view, including both successes and challenges. An absence of such balance may indicate a lack of sincerity. Furthermore, qualitative studies have revealed discrepancies between what companies state in their ESG reports and external evaluations of their performance. Parguel et al., (2011), for instance, examined cases in which a company presented itself as a climate leader in its report, while independent ESG ratings identified it as a major emissions producer, which revealed a credibility gap.

In recent years, natural language processing (NLP) and artificial intelligence (AI) have greatly enhanced our ability to systematically analyse ESG reports. Rather than treating sustainability disclosures as dense, unreadable brochures, researchers can now convert textual content into quantifiable data (Ong et al., 2025). This enables the measurement of various linguistic features, such as tone, readability, specificity, and even levels of optimism or candor. For example, Kang and Kim (2022) used text mining techniques to corporate sustainability reports and found that while some firms employ detailed and varied language aligned with substantive ESG themes, others rely heavily on repetitive and generic expressions. This suggests differing levels of disclosure sincerity. Such studies highlight the wide variation in the quality of ESG reporting (in terms of concreteness and balance) across firms, and the correlation between certain textual traits and more credible sustainability performance.

In order to detect instances of greenwashing, researchers have identified specific linguistic indicators that suggest a report may be more concerned with managing perceptions than with ensuring accountability. Prominent indicators identified in the literature include the use of excessively positive tone, vague buzzwords, hedging language, lack of specificity, and an absence of candor or balance in reporting.

An overly optimistic or congratulatory tone in a report can be a warning sign. Li et al. (2024) provide empirical evidence that an 'abnormally positive' sentiment in ESG disclosures indicates a higher ESG risk for the firm. Therefore, managers overstating good news in sustainability reports may be doing so to mislead, and such firms often encounter ESG problems later on. Similar patterns have been observed in financial communication, where companies engaged in impression management tend to favour highly positive language while avoiding negative terms. Melloni et al. (2016) found that a more negative tone in integrated reports (acknowledging problems) tended to reflect more serious sustainability efforts. By contrast, disclosures that only refer to “excellent progress” without mentioning any “challenges” may reflect a lack of honesty about environmental impact. Truly transparent reports usually contain at least some neutral or negative language when describing ongoing challenges. The absence of any negative wording, together with overly positive phrasing, raises questions about the credibility of a report.

Another key indicator of greenwashing is the overuse of ESG-related buzzwords and jargon. As discussed in Section 2.1, the use of vague ESG buzzwords is a widespread greenwashing tactic. Reports containing terms such as 'eco-friendly', 'sustainable', 'green' and 'innovative impact' without concrete context or definitions are not uncommon. Such vague language can create the impression of action while actually saying very little. Delmas and Burbano (2011) observe that confusing statements (e.g. labelling a product 'all-natural' despite the term having no clear definition or supporting evidence) exploit stakeholders' positive associations while providing no concrete information. A recent analysis by *Bridge* (2024) highlights that vague or generic ESG claims, particularly those full of feel-good buzzwords, can damage credibility, as investors are becoming better at recognising and rejecting empty promises.

Another sign is the presence of abundant hedging language, such as words that express uncertainty, aspiration or non-commitment. Companies engaging in greenwashing may fill their disclosures with qualifiers such as 'aim to', 'commit to exploring', 'strive to', or modal verbs like 'may', 'might', or 'could'. While some use of such language is normal, as future plans are naturally uncertain, overuse suggests that the firm is avoiding making commitments. By relying on hedging, firms can create the appearance of action or ambition without making verifiable promises. For instance, Vinella et al. (2024) identify vague and non-committal phrasing as key markers of greenwashing in corporate sustainability reporting. Similarly, another study found that large language models (LLMs) could effectively detect greenwashing by analysing the prevalence of hedging language in climate-related disclosures (Chuang et al., 2025).

Closely related is the issue of specificity. Those making greenwashed disclosures often avoid concrete metrics or details. For instance, a report might state that 'significant improvements in waste reduction' have been made, but not quantify anything or explain how this has been achieved. In contrast, a credible report would provide specific figures (e.g. 'a 15% reduction in harmful waste this year, against a target of 20%'). A recent study by Hassan (2024) found that firms relying on vague sustainability language without supporting data are significantly more likely to engage in superficial ESG practices, rather than implementing meaningful change. This highlights the importance of specificity: the more concrete, numeric and outcome-based a disclosure is, the less room there is for greenwashing.

Finally, the presence (or absence) of candor is widely regarded as an indicator of report credibility. As mentioned previously, the omission of failures, challenges, or past ESG controversies in a report may signal attempts to manage impressions rather than inform stakeholders (Hahn & Lülfs, 2014). Truly transparent firms will acknowledge areas where targets have been missed or challenges persist. By contrast, greenwashing firms will either ignore negative aspects entirely or bury them in unclear language. Researchers have proposed metrics for candor, such as the proportion of sentences discussing negative performance or the presence of self-criticism. Higher candor scores (openness about shortcomings) are generally associated with more trustworthy ESG practices, while lower scores suggest the report functions more as a promotional tool than a disclosure instrument. One way to assess candor is by identifying keywords that signal problems, controversy, or past criticism. If no such language is present, the report's credibility should be viewed with caution, as it may be presenting an overly embellished and unrealistic narrative.

More broadly, the emergence of sophisticated natural language processing (NLP) techniques, particularly based on transformer models, has enabled new methods for detecting greenwashing. BERT-based models are widely used in ESG analysis for their ability to capture context. For example, Kim et al. (2017) used a domain-specific ESG-BERT model alongside machine learning classifiers to analyse the sustainability reports of Korean companies and achieved significant success in identifying greenwashing cases. These AI tools offer two major benefits: they can analyse large volumes of text rapidly and detect subtle linguistic patterns.

However, challenges remain, particularly in interpretability. While sophisticated models may flag a report as greenwashing with high confidence, explaining why can be difficult (i.e. which phrases or inconsistencies led the model to make this decision). Research in explainable AI (XAI) addresses this by highlighting words or features that most influenced the classification. Legal-Pythia's platform, for instance, combines complex data analysis with clear explanations of what drove the classification (*Ocean Business B2match*, 2025). Another issue is the need for labelled data to train these systems. Some studies rely on known cases (e.g. scandals or NGO reports) to classify certain disclosures as greenwashing and then train models based on this information. Vinella et al. (2024), for example, fine-tuned ClimateBERT on such data, achieving 86.34% accuracy. Still, this approach is limited, as many incidents remain unreported.

2.4. Large Language Models (LLMs) in ESG and Greenwashing Analysis

The emergence of large language models such as OpenAI's GPT-3/4 (ChatGPT) has introduced new opportunities for analysing ESG communications. LLMs are trained on vast amounts of data and can understand and generate human-like text (Wei et al., 2022). This enables them to interpret context, nuances and implied meanings in a way that earlier algorithms could not. This capability is particularly relevant to ESG and greenwashing analysis, as these areas often depend on the subtleties of language and tone. For example, Zou et al. (2023) introduced ESGReveal, a framework that uses LLMs combined with Retrieval-Augmented Generation techniques to systematically extract and analyze ESG data from corporate reports. The study demonstrated that their framework achieved accuracy rates of 76.9% in data extraction and 83.7% in disclosure analysis, outperforming baseline models.

One area in which LLMs are being applied is the summarisation and interpretation of ESG reports. Analysts can save considerable time by using a model like ChatGPT to rapidly produce a summary of a long sustainability report, which highlights the main points, environmental targets, social programs and governance changes. This is exemplified in the DeepGreen framework, which uses a dual-layer LLM architecture to extract and assess ESG-related content from financial reports (Xu et al., 2025). Furthermore, LLMs can be prompted to answer questions about a report's content. For instance, an investor or researcher could ask ChatGPT, “Does Company X’s sustainability report provide evidence of progress on its climate goals, or does it employ vague language?”. The model can analyse the text and provide an answer that highlights whether or not specifics are provided. Early experiments show that ChatGPT can often identify when text contains mostly generic statements versus concrete facts (Dong et al., 2024). This suggests that LLMs could be used as a first line of defence against greenwashing by effectively asking the AI, 'Is this report credible, or does it sound like greenwashing?'

LLMs are also being explored for cross-document consistency checks. Since they have extensive knowledge, it is possible to provide the model with relevant information (Leippold et al., 2024). One could, perhaps, submit both the company's sustainability report and some external data, such as news articles or NGO reports, and then ask the model to compare them. It might then point out discrepancies, such as the company claiming a perfect safety record while media reports document two accidents. Surana et al. (2022) have already demonstrated similar capabilities in legal contexts, where models can flag inconsistent statements across documents This task requires an understanding of context and reasoning based on language meaning. In ESG analysis, this functionality could help identify greenwashing by detecting discrepancies that manual review would struggle to uncover.

Furthermore, researchers are starting to fine-tune LLMs for classification tasks related to ESG, with the aim of improving their performance and accuracy in these areas. A recent study by Huang et al. (2024) tested GPT-4’s ability to classify sentences from sustainability reports as either *“concrete action”* or *“vague promise”*. The model achieved high accuracy, which demonstrates its potential to flag greenwashing at the sentence level.

However, limitations remain. These models can sometimes 'hallucinate', which means they may generate an answer that sounds plausible but is not actually supported by the text (Banerjee et al., 2024). This is problematic if investors were to rely on them without verification. There is also the question of bias: ChatGPT's training data may contain certain perspectives on corporate behaviour that influence its judgements. For example, some industries might be viewed with more scepticism due to the content that was seen during training. Researchers must therefore check the outputs of LLM technology against established facts and be careful when interpreting them. Despite these challenges, preliminary assessments suggest that generative AI, such as ChatGPT, can greatly simplify the analysis of sustainability reports (Dong et al., 2024).

2.5. Conclusions

These findings are of great importance for researchers and practitioners. As the literature shows, large language models (LLMs), such as ChatGPT, are being used more and more to summarise, interpret and cross-check ESG disclosures against external data sources (Rao & Srinivasa, 2022; Bronzini et al., 2023). The ability to detect inconsistencies and derive meaningful insights in real time opens up new possibilities for identifying greenwashing. Although this application is still in its early stages, it is growing rapidly (Leippold et al., 2024). If sophisticated AI models can reliably identify vague, exaggerated, or contradictory sustainability claims, investors could be notified before such claims distort market valuations. This would help markets evaluate the credibility of ESG disclosures more efficiently and would encourage firms to report more honestly and substantively.

3. Methodology

3.1. Research Question and Hypotheses

This chapter details the empirical strategy developed to address the central research question posed in this thesis: *How can artificial intelligence-driven language analysis models enhance the detection of greenwashing tactics in corporate ESG disclosures, and what influence does this detection have on the stock market?*

To investigate this, we implement a two-step methodology combining textual analysis and an event study approach. First, we calculate a Greenwashing Score (GWS) for each ESG report, capturing linguistic traits linked to symbolic or insincere communication. Inspired by prior work but newly designed here, the rule-based GWS includes five dimensions: buzzword density, specificity, sentiment tone, hedging, and candor. Each is computed using NLP tools and normalised by industry. To complement this score, an alternative AI-driven model based on a large language model (GPT-4) was trained to replicate human interpretation of ESG discourse, providing an additional layer of analysis.

Second, we conduct an event study to examine how markets respond to ESG reports. Using cumulative abnormal returns (CARs) as a proxy for investor reaction, we test whether firms with higher GWS experience less favourable market responses. A regression framework with firm- and industry-level controls is used to assess this relationship.

The sections that follow outline the dataset, the construction of the Greenwashing Score, and the event study methodology used to evaluate its financial implications. From these objectives, we derive two key hypotheses to be tested:

- **H1 (Greenwashing Prevalence):** The level of greenwashing in ESG disclosures can be quantified using AI-based textual analysis. In other words, we expect the Greenwashing Score to reveal meaningful differences in firms' communication styles, with some reports exhibiting significantly more 'greenwashing' characteristics (e.g. buzzwords and vagueness) than others. This hypothesis lies at the heart of the development of the GWS metric, which assumes that linguistic indicators can successfully indicate when a report is embellished or inaccurate.
- **H2 (Market Impact):** Firms with higher GWS, which suggests they may be more likely to greenwash in their ESG reports, are expected to see negative impacts on their CARs around the time of report release. This would imply that investors either penalise insincere disclosures or reward them less. If investors can discern highly optimistic or vague ESG reporting, we would expect to see a negative correlation between GWS and CAR (i.e. higher greenwashing would be associated with lower or negative abnormal returns).

3.2. Data and Sample

This study relies on a dataset of over 1,500 annual ESG or sustainability reports from publicly traded companies, covering multiple industries and years. Each report is linked to firm-level financial data and daily stock returns, enabling a broad cross-sectional and temporal analysis.

The sample includes firms of varying sizes and sectors, from energy and materials to technology and finance, which allows us to account for sector-specific communication styles and investor expectations. Prior to analysis, all report texts were pre-processed to enable reliable textual analysis. This included cleaning procedures such as removing HTML tags and non-content elements, and also standardising whitespace and case. All reports were converted into machine-readable format to enable the extraction of linguistic features using natural language processing tools. The analysis was conducted in R, primarily with the tidytext, stringr, and textdata libraries. This processing pipeline forms the foundation for constructing the Greenwashing Score, presented in the following section.

3.3. Greenwashing Score (GWS) Construction

In order to detect greenwashing in ESG disclosures, we have developed a Greenwashing Score composite for each report. It quantitatively assesses five core linguistic components of the report's content, all of which are based on previous research into impression management, the quality of corporate disclosure, and sustainability communications. These five components capture known textual indicators of potential greenwashing (see Appendix 1).

3.3.1. *ESG Buzzword Density*

This component measures the extent to which a report is filled with popular ESG buzzwords and jargon. We put together a dictionary of common sustainability buzzwords and phrases (e.g. 'sustainable', 'green', 'eco-friendly', 'net zero', 'carbon neutral', etc.) and counted how often they appeared in each report. We then calculate the frequency of these words per 1,000 words to normalise for report length. The use of a high number of ESG buzzwords may indicate symbolic disclosure, or that attention is being given to form rather than substance.

Prior research shows that companies may use a lot of ESG buzzwords to mislead or impress stakeholders without providing real substance (Vereb, 2024). Studies in accounting and management communication have similarly linked the frequent use of positive ESG language to 'cheap talk', especially when such claims lack supporting evidence (Merkl-Davies & Brennan, 2008). Moreover, Boiral et al. (2021) describe this as ‘symbolic management’, the use of broad, feel-good terminology disconnected from actual performance. In the same vein, Cho et al. (2015) argue that excessive ESG jargon can create a false sense of transparency, masking the absence of meaningful disclosure.

Complementing this view, a 2024 study finds that lower readability in ESG reports is significantly associated with higher levels of greenwashing (Hu et al., 2024). They argue that complex language and excessive rhetoric may serve to obscure the absence of concrete actions. This reinforces the inclusion of buzzword density as a key indicator, as firms that rely heavily on vague or fashionable ESG terminology, especially without sector-specific context, may be engaging in misleading disclosure practices.

By benchmarking a firm's use of buzzwords against industry norms, we can more easily identify disclosures that rely disproportionately on language rather than facts. This highlights potential greenwashing based on language choice rather than content alone. If the score is significantly higher than the industry average, this suggests that the company may be placing too much importance on popular terms.

3.3.2. Lack of Specificity (Vagueness)

This component assesses how concrete and measurable the claims in an ESG report are. We determine the level of detail in the text, looking for quantifiable and verifiable information such as numerical figures, specific targets or statistical disclosures related to ESG performance. Statements such as 'CO_2 emissions were reduced by 15% in 2024' are considered highly specific, whereas expressions such as 'significant reduction' or 'we are improving our sustainability' lack concrete metrics and are considered vague.

We calculate a specificity ratio based on numeric performance indicators and targets. We measure this as a specificity ratio and this is calculated per 1,000 words of text. Reports containing few numerical details and relying heavily on qualitative or generic statements receive lower specificity scores, which indicates a higher degree of vagueness. This aligns with previous research findings: a low level of measurable disclosure is often considered a sign of greenwashing (Melloni et al., 2017).

Greenwashing often takes the form of broad, unverifiable promises rather than concrete data (Team, 2025). Stakeholders are increasingly relying on quantifiable targets to assess the credibility and transparency of corporate sustainability reporting. A recent study by *The Choice* (2025) on AI-based ESG fraud detection identified vague language and the absence of measurable targets as internal red flags. In line with this, our specificity metric reflects the idea that transparent ESG reports should provide concrete figures, such as emissions, diversity percentages and investment amounts, rather than non-specific slogans.

To construct a robust and comparable measure, we normalise the raw count of numeric disclosures per 1,000 words and apply a logarithmic transformation ($\log(1 + x)$) to reduce the skewness caused by reports with extreme values. This gives us a Specificity Score for each report, where higher values indicate more specific, quantified content. The resulting Specificity Score is then inverted (1 – specificity), in order to transform it into a vagueness indicator. In short, reports with minimal quantitative content contribute positively to the greenwashing risk index.

3.3.3. Sentiment Tone

The tone of language used in an ESG report, particularly its sentiment polarity, is another important indicator. Research in impression management suggests that companies may emphasise positive language strategically to portray themselves in a favourable light, while downplaying risks or negative outcomes (*KPMG*, 2022). To quantify this, we use the Loughran & Mcdonald (2011) financial sentiment lexicon which is tailored for corporate disclosures and avoids misclassifying financial terms (Dame, n.d.).

For each report, we calculate the number of words categorised as positive (e.g. 'progress', 'achievement', 'beneficial') or negative ('concern', 'failure', 'risk') according to the dictionary. These counts are then used to calculate a net sentiment score (see Appendix 1). Higher scores reflect a more optimistic tone. As noted in the literature, excessive positivity, especially when abnormal relative to peers, can indicate impression management (Li et al., 2024).

As tone naturally varies across industries (for example, tech companies often adopt a more optimistic tone than firms in the manufacturing sector), we normalise sentiment scores at the industry level. For example, a fossil fuel company with a sentiment score that far exceeds the average tone of its peers may be identified as potentially engaging in greenwashing through optimism bias.

In this framework, a higher Sentiment Tone Score contributes positively to the overall Greenwashing Score. This approach aligns with emerging evidence that investors react not only to what ESG information is disclosed, but also to how it is communicated.

3.3.4. Hedging Language

This component captures the use of language that signals uncertainty or avoids commitment, such as "may", "might", "aim to", "approximately", or "aspire to". In ESG reporting, excessive hedging can suggest reluctance to be held accountable and may obscure what has actually been achieved. Companies engaged in greenwashing often rely on aspirational or conditional wording to avoid making firm promises.

We quantify hedging by counting the frequency of terms from a custom list of common hedge words, based on prior literature. This count is normalised per 1,000 words and log-transformed to account for skewness in the distribution.

Reports containing an unusually high number of hedging terms are seen as potentially less credible, as such language may make it difficult to hold the firm accountable. Kotte et al. (2019) suggest that companies sometimes deliberately use vague and evasive language in their ESG disclosures in order to avoid accountability or legal liability. This has the reverse effect of making the disclosures less concrete and informative. Theil and Stuckenschmidt (2020) similarly link excessive hedging to greater uncertainty in communications, while Dzieliński et al. (2017) find that executives' modal language is predictive of volatility in stock prices and analyst forecasts.

Although some hedging is normal, a report filled with qualifiers but lacking concrete performance statements may indicate disengagement or strategic obfuscation. In our model, higher hedging scores contribute to the overall greenwashing risk, especially when combined with low specificity or high vagueness.

3.3.5. *Candor (Transparency about Shortcomings)*

The final component assesses how forthcoming the report is about negative aspects or shortcomings regarding the company's ESG performance. Melloni et al. (2017) use the term 'balance' in reporting to refer to the inclusion of negative information. They find that reports from better-performing firms tend to be more balanced, whereas poor performers engage in impression management to avoid negative disclosure.

To assess candor, we created a list of self-critical phrases reflecting the admission ESG-related shortcomings. Based on sustainability reporting guidelines and previous research into negative disclosures, the list includes phrases such as 'did not meet our target', 'fell short', 'behind schedule', 'room for improvement', 'challenges remain', 'missed our objective', 'needs improvement' and 'lessons learned'. The Candor Score reflects the frequency of these terms, normalised by report length.

Because a lack of candor is what increases greenwashing risk in our framework, we invert the score. So, Candor Risk is equal to 1 - Candor Score. A company with zero admissions of difficulty receives a candor risk score close to 1 (high risk), whereas a report that is self-critical scores closer to 0 (lower risk).

This concept aligns with stakeholder expectations: investors and consumers appreciate honesty and are more likely to trust reports that do not hide imperfection (Team, 2025). As one guideline suggests, companies should *"prioritise transparency over perfection"* and that stakeholders *"value honesty, even if a company is still in the process of improving its sustainability performance"* (Team, 2025).

However, we acknowledge that the absence of unfavourable information is not necessarily misleading. A firm that has had a genuinely successful year in terms of ESG factors may have little negative news to report. To reflect this nuance, we give less weight to the candor component in the Greenwashing Score composite. While a lack of candor can be seen as a warning sign, it should be interpreted with prudence, particularly in cases where strong performance could reasonably result in fewer admissions of failure.

3.4. Normalization and Composite Scoring

Each of the five components produces a numeric score for a report, with higher or lower values indicating more or less of that characteristic. However, these raw scores are on different scales and distributions. To combine them into a single Greenwashing Score, we apply several normalisation steps.

3.4.1. Industry Adjustment

To account for systematic differences in communication styles across industries, we standardise certain components of the Greenwashing Score using industry-relative z-scores. Specifically, we apply this adjustment to buzzword density and sentiment tone, as what qualifies as "high" or "excessive" can vary substantially between sectors. For instance, tech firms may naturally use more optimistic or buzzword-heavy language than firms in oil and gas or manufacturing.

For each report, we calculate a z-score for these components by subtracting the industry mean and dividing by the industry standard deviation. This transformation tells us how far a company's score deviates from the typical level observed in its sector. A high positive z-score indicates unusually dense buzzword usage or sentiment tone, even after accounting for industry norms. This adjustment ensures that our Greenwashing Score flags abnormal disclosure behaviour relative to peers, without penalising sectors for typical communication patterns.

3.4.2. Scaling to [0,1]

After industry adjustment, we rescale each component to the range 0–1 to ensure comparability. In practice, we apply min–max normalisation to each metric within the full sample or within the industry, as appropriate. The highest observed value is scaled to 1 (high presence of the trait) and the lowest to 0, with the others scaled proportionally in between. To ensure consistency across metrics, all scores are oriented so that higher values reflect greater greenwashing risk. For others where low values imply risk (specificity and candor), we invert the scale, for instance, low specificity becomes high vagueness, and low candor becomes high opacity. At the end of this step, we have five normalised component scores for each report, all of which are between 0 and 1 and aligned in direction.

3.4.3. Weighting and Aggregation

The final Greenwashing Score is a weighted average of the five component scores. Rather than averaging them, we assign differential weights to reflect the relative importance of each indicator type, informed by theory and practical considerations. We determined the weighting scheme based on our judgement of which factors most strongly signal greenwashing risk, while also considering the reliability of each measure. The weights, which sum to 5.0 for ease of interpretation, are distributed as follows: buzzword density is weighted at 1.5 (30%), lack of specificity at 1.2 (24%), sentiment tone at 1.0 (20%), hedging language at 0.8 (16%), and lack of candor at 0.5 (10%).

We assigned the highest weight to buzzword usage because relying heavily on ESG jargon without having any substantive content is often the most visible sign of greenwashing, and it appeared frequently in the literature as a red flag (Boiral et al., 2021; Merkl-Davies & Brennan, 2008). The specificity/vagueness component was given the second-highest weighting, as vague claims are a classic feature of questionable communication. Sentiment tone was given a weighting of 1.0 as an excessively positive tone is important, but we did not want to overstate its importance given that some firms naturally have a positive style. Hedging was assigned a lower weight of 0.8, reflecting the fact that, while hedging can indicate insincerity, moderate hesitant wording may simply be due to uncertainty in forward-looking statements (Theil & Stuckenschmidt, 2020). Finally, candor/transparency was assigned the lowest weight, as a lack of negative disclosures can sometimes indicate genuinely good performance rather than intentional omission. Thus, although this indicator is meaningful, it should not have a significant impact on the composite score.

After applying these weights, the component scores are summed to generate the composite greenwashing score, which ranges from approximately 0 (low risk) to 1 (high risk). This continuous metric reflects the degree of greenwashing inferred from each report's language and serves as the key independent variable in our analysis of market reactions.

3.4.4. Validation and Robustness of the Greenwashing Score

To ensure the reliability and conceptual validity of the Greenwashing Score, we conducted several validation checks on its five underlying components. These aimed to verify that each captures distinct and meaningful aspects of greenwashing, rather than overlapping concepts or random variation.

3.4.4.1. Component Independence

Firstly, we examined the intercorrelations between the five linguistic indicators used to construct the GWS. Pairwise Pearson correlations revealed low to moderate relationships across all variables, with absolute correlation coefficients below 0.45. This suggests that each component captures a distinct aspect of ESG communication. For example, buzzword usage and hedging showed virtually no correlation ($r = 0.03$), while the strongest relationship, between vagueness and hedging, remained moderate ($r = 0.45$). These results confirm that the GWS integrates diverse, non-redundant dimensions and is not driven by any single factor (see Appendix 2 for the full correlation matrix).

3.4.4.2. Contribution of Each Component to the GWS

To assess the contribution of each component to the overall Greenwashing Score, we computed Pearson correlations between each indicator and the composite GWS. As expected, all components were positively associated with the GWS, reflecting their inclusion in its calculation. The strongest relationship was found with buzzword density ($r = 0.84$), followed by sentiment tone ($r = 0.46$), vagueness ($r = 0.37$), hedging ($r = 0.36$) and lack of candor ($r = 0.11$). These results are consistent with the theoretical weights assigned to each component, and confirm that the composite score provides a multidimensional assessment of greenwashing risk (see Appendix 3 for full correlation values).

3.4.4.3. Multicollinearity Check

To test for multicollinearity among the components when used together in a predictive framework, we ran a linear regression analysis, treating the GWS as the dependent variable and the five components as the independent variables. We then computed the variance inflation factor (VIF) for each predictor. All VIF values were close to 1, ranging from 1.04 to 1.36, well below the commonly accepted threshold of 5 (see Appendix 4). This confirms that the components do not introduce redundancy into the model and can be interpreted as statistically independent inputs.

3.5. Collaboration with an AI Co-Analyst for Score Validation

The core aim of this thesis is to develop a composite Greenwashing Score using AI-driven linguistic analysis of ESG reports. The score integrates five textual components (as mentioned in Section 3.3) in order to quantify a firm's tendency towards "greenwashing". Each element captures a different aspect of potential misrepresentation, from excessive jargon to a lack of concrete detail.

To strengthen the validity of the index, we conducted a cross-validation exercise using OpenAI's GPT-4, a cutting-edge large language model (LLM), as an independent evaluator. We provided ChatGPT with a detailed prompt (see Appendix 5) describing the five components and their general intent. Importantly, we employed a chain-of-thought prompting technique to encourage the model to reason through each dimension step by step (Wei et al., 2023). This approach produces intermediate justifications before delivering a final judgement. Not only does this method improve interpretability, it also mirrors the reflective reasoning process of a human analyst.

While the evaluation criteria were intentionally aligned with our own to ensure comparability, GPT-4 was given full freedom to interpret, weight, and transform the components based on its own logic. This was not intended to constrain the model's reasoning, but rather to ensure direct comparability between the two systems. This setup allowed us to assess whether the AI, using a distinct reasoning process, would prioritise similar features and detect comparable patterns of symbolic ESG communication.

3.5.1. Inferred Weighting Structure (GWS_gpt)

Although the model was not provided with a formula, we were able to infer its scoring behaviour based on its outputs and justifications that GPT-4 generated, we derived an implicit weighting scheme that reflects how the model combined the five components into a composite greenwashing score (GWS_gpt). The resulting formula is as follows:

$$GWS_{gpt} = 0.30 \times (1 - Specificity) + 0.20 \times Tone + 0.20 \times Hedging + 0.15 \times Buzzwords + 0.15 \times (1 - Candor)$$

These weights were inferred from GPT-4's explanations and reflect its own reasoning, not predefined instructions. The rationale for each is outlined below.

3.5.2. Rationale for GPT-4's Weights

When asked to evaluate ESG disclosures without a preset formula, GPT-4 implicitly adopted a weighting structure that prioritised a lack of specificity (i.e. vagueness) over all other dimensions. This was followed by a moderate emphasis on sentiment tone and hedging, and a more careful consideration of buzzword use and a lack of candor (see Appendix 6). The logic behind these inferred weights is detailed below.

3.5.2.1. Specificity score — Weight: 0.30

GPT-4 consistently penalised reports missing concrete and verifiable ESG metrics. Statements such as 'working towards a greener future' were considered meaningless unless supported by specific targets, such as 'reducing emissions by 18% in 2023'. This mirrors the concept of 'symbolic management' in the literature, whereby firms make grand claims with little substance (Boiral, 2013; Merkl-Davies & Brennan, 2008). Given its extensive training data, which includes regulatory texts, corporate filings and financial journalism, it is likely that GPT-4 "understands" that specificity is a key indicator of accountability. The component's weighting reflects this insight.

3.5.2.2. Sentiment Tone — Weight: 0.20

GPT-4 often detected overly optimistic or triumphant language as strategic impression management, especially when it was imbalanced. Phrases such as 'we lead the sustainability transition' raised concerns unless they were balanced by a mention of risks or limitations. While some positivity is expected in ESG reporting, GPT-4 adopted a rather sceptical attitude when the tone lacked nuance, which aligns with the findings of Cho et al. (2015). Therefore, tone was given significant weight, but not dominant weight.

3.5.2.3. Hedging Language — Weight: 0.20

Soft commitments were identified as phrases such as "we aim to", "may" or "intend to". GPT-4 raised alerts for reports where such language dominated without being paired with concrete actions. This reflects patterns in executive communication where hedging is employed to manage expectations or avoid liability (Yang, 2019). Nevertheless, GPT-4 appeared to recognise that some hedging is normal in forward-looking disclosures, which is why it assigned this dimension the same moderate weight as sentiment tone.

3.5.2.4. Buzzword Usage — Weight: 0.15

GPT-4 recognised buzzwords such as "green", "eco-friendly" and "impact-driven", but evaluated them in context. Reports that combined buzzwords with clear metrics were often considered more credible. Research suggests that the use of ESG buzzwords does not in itself indicate dishonesty; rather, such language is better understood as a potential symptom of greenwashing, not definitive proof of it (Vereb, 2024). GPT-4 appeared to adopt this approach by weighing them as a secondary indicator rather than a critical one.

3.5.2.5. Candor — Weight: 0.15

GPT-4 acknowledged that the omission of failures or missed targets could indicate green silence, but also recognised that such omissions may stem from legitimate reasons, such as strong ESG performance or legal prudence. Given this ambiguity, it treated lack of candor as a contextual indicator, rather than a primary red flag. This approach is consistent with the findings of Melloni et al. (2017), who point out that the absence of negative information does not necessarily indicate dishonesty.

3.5.3. Scoring Comparison and Interpretation

We compared GPT-4's greenwashing assessments with our rule-based Greenwashing Score (GWS), both built on the same five linguistic components, to assess alignment between the models. This aimed to validate our approach and explore how GPT-4's human-like, contextual reasoning compares with our structured, lexicon-based method.

The Pearson correlation between our GWS and GPT-4's score was r = 0.276, which indicates a moderate positive relationship (see Appendix 7). While this confirms that the two models identify similar patterns of greenwashing risk to some extent, it also reflects their fundamentally different evaluation logics. Our GWS is calculated from predefined weights and normalised indicators. GPT-4 produced its score through contextual interpretation. When prompted, it was able to explain its reasoning and the relative importance of each component, allowing us to reconstruct an explicit weighting scheme.

A comparison of component scores reveals deeper contrasts (see Appendix 8). Our model places greater emphasis on buzzword usage and lack of candor, which is consistent with its focus on specific lexical cues. GPT-4, however, downplays buzzwords when embedded in goal-oriented discourse and is more sensitive to vague commitments and promotional tone, particularly in the hedging and sentiment dimensions. For example, the median difference in hedging scores exceeds 0.5.

The strongest alignment between the two models occurs in the 'Specificity' dimension, which GPT-4 also identified as the most critical indicator of greenwashing. This reinforces findings from sustainability reporting literature on the importance of measurable ESG commitments. More broadly, GPT-4 behaves like a qualitative auditor: it interprets not only what is said, but how and why it is framed. Together, these findings suggest that rule-based and LLM-based assessments are complementary, rather than contradictory. While our model offers transparency, replicability, and comparability, GPT-4 provides interpretive depth, making it more sensitive to narrative-driven signals.

3.6. Event Study: Stock Market Reaction to ESG Reports

The second part of the methodology, which uses the Greenwashing Score to assess each ESG report, investigates whether the stock market reacts differently depending on the perceived credibility of ESG communication. To explore this, we use the event study methodology (MacKinlay, 1997), which is widely employed in finance to evaluate investor responses to new information. Specifically, we examine how company stock prices behave around the time their sustainability reports are released.

Our main indicator of investor response is the Cumulative Abnormal Return (CAR), which captures a stock's performance relative to expected market returns during a defined event window (Müller, 2024). CARs help to isolate the effect of the ESG disclosure itself by comparing actual stock movements with those that would typically be expected in the absence of new information. A positive CAR suggests that the market responded favourably to the disclosure, which may indicate credibility or strong performance. Conversely, a negative CAR may reflect disappointment, scepticism, or perceived insincerity.

To capture market reactions across different time horizons, we use three CAR windows: CAR[0,2], which reflects short-term responses over three trading days; CAR[0,5], representing a one-week window; and CAR[0,60], which captures medium-term effects over approximately three months. The next step is to link these CARs to the Greenwashing Score. We test whether perceived greenwashing affects investor confidence by regressing the various CAR measures on the GWS (and relevant control variables). If firms with higher GWS values consistently exhibit lower CARs, this would suggest that markets are sensitive to the quality and sincerity of ESG disclosures, and that greenwashing could have real financial implications.

3.7. Regression Analysis: Linking Greenwashing Scores to Stock Market Reactions

To deepen the analysis, we run a set of firm-level regressions linking the Greenwashing Score to investor behaviour. The objective is to examine whether reports flagged as more "greenwashy", based on their linguistic features, are systematically associated with weaker market reactions. We estimate four distinct models, each capturing a different dimension of this relationship, from the direct impact of greenwashing signals to potential interaction effects. This approach allows us to explore not just whether greenwashing matters to investors, but how and under what conditions it influences their responses.

3.7.1. Dependent Variables

The dependent variable across all models is the Cumulative Abnormal Return for firm *i* around the ESG disclosure date *t*. and more precisely CAR2, CAR5 and CAR60 (see Section 3.6). These allow us to capture both immediate and delayed market responses. We use CAR as the dependent variable because our central research objective is to assess how investors react to perceived greenwashing in ESG communications and this is the most appropriate financial outcome to detect valuation effects linked to language-driven credibility signals.

3.7.2. Independent Variables

The main independent variable across all regression models is the Greenwashing Score, constructed using two distinct AI-based approaches (see Sections 3.3 and 3.5). The first method employs a rule-based natural language processing (NLP) framework to produce a composite score (greenwashing_score) based on five linguistic dimensions. The second approach uses GPT-4 to evaluate the same five components and generate an alternative score (GWS_gpt) based on its own weighting logic. In regression Models 1 and 2, we use the composite scores as predictors to assess their overall association with market reactions. In Models 3 and 4, we disaggregate the score into its five underlying components, treating each linguistic feature as a separate explanatory variable. This allows for a more detailed analysis of the specific disclosure traits that influence market responses.

3.7.3. Control Variables

To control for factors that might influence market reactions independently of ESG credibility, we include a set of control variables. Firm size is measured as the natural logarithm of (1 + total assets) to account for differences in visibility, market power, and investor attention. Profitability is captured by return on assets (ROA), since more profitable firms may elicit different investor responses. We also control for stock momentum, to account for recent performance trends that could influence CARs regardless of ESG behaviour. Finally, we include each firm's external ESG rating, which reflects its assessed sustainability performance and helps isolate the effect of linguistic credibility from actual ESG outcomes.

Additionally, we include industry and year fixed effects to account for unobserved heterogeneity. Industry fixed effects control for sector-specific norms in ESG reporting and typical investor expectations. Year fixed effects capture the macroeconomic and regulatory conditions that may influence ESG disclosures or market responses in a given year. These fixed effects ensure that our estimated coefficients reflect variation within industries and years. So, it allows us to isolate the relationship between greenwashing and CARs from broader trends.

3.8. Regression Models

3.8.1. Model 1 (M1) – ESG Credibility and Market Reactions

The first model provides the foundational framework for examining whether greenwashing, as detected by AI-based textual analysis, has a measurable effect on stock market reactions. Specifically, this regression model tests the relationship between the greenwashing score (both NLP & LLM based models) and the firm's cumulative abnormal returns over multiple event windows: two days, five days, and sixty days around the release of ESG reports. The regression takes the following general form:

$$CAR_{i,t} = \beta_0 + \beta_1 GWS_{i,t} + \beta_2 log(1 + SIZE_{i,t}) + \beta_3 ROA_{i,t} + \beta_4 MOMENTUM_{i,t} + \beta_5 ESGscore_{i,t} + \gamma_{Industry} + \delta_{Year} + \varepsilon_{i,t}$$

This model directly tests Hypothesis 2 (see Section 3.1), which suggests that greater amounts of greenwashing in ESG reports are linked to negative market responses. A statistically significant negative coefficient on either greenwashing measure would suggest that investors react unfavourably to perceived dishonest or exaggerated ESG reports. This model also indirectly supports Hypothesis 1 by suggesting that the GWS metrics capture meaningful variation in the linguistic quality and credibility of ESG communication across firms if they consistently explain market behaviour.

3.8.2. Model 2 (M2) – Interaction with ESG Performance: Contextualizing Greenwashing

This model builds on the original regression analysis by introducing an interaction term between the greenwashing score and the firm's ESG performance rating. It aims to explore whether the market's reaction to perceived greenwashing depends on the firm's underlying ESG quality. In other words, it investigates whether investors react more negatively to greenwashing when it is associated with firms that are otherwise perceived as high ESG performers. The model can be expressed as:

$$CAR_{i,t} = \beta_0 + \beta_1 (GWS \times ESGscore)_{i,t} + \beta_2 log(1 + SIZE_{i,t}) + \beta_3 ROA_{i,t} + \beta_4 MOMENTUM_{i,t} + \gamma_{Industry} + \delta_{Year} + \varepsilon_{i,t}$$

This interaction-based model adds depth to Hypothesis 2 by suggesting that market participants do not evaluate greenwashing in isolation. Rather, investors may be particularly sensitive to gaps between a firm's self-proclaimed excellence in environmental, social and governance matters and the verbal cues in its reporting. The presence of a negative and significant interaction term would suggest that firms are penalised more severely by the market when strong ESG ratings are paired with communication traits typically associated with greenwashing.

3.8.3. Model (M3) – Disaggregated Linguistic Features: Testing the Components of Greenwashing

Model 3 decomposes the composite Greenwashing Score into its five components: sentiment tone, lack of specificity, buzzword density, hedging, and candor. These dimensions are entered individually into the regression to examine their distinct associations with abnormal returns, while controlling for firm characteristics, ESG scores, and fixed effects. The formula for the regression is as follows:

$$CAR_{i,t} = \beta_0 + \beta_1 Sentiment_{i,t} + \beta_2 Specificity_{i,t} + \beta_3 Buzzwords_{it} + \beta_4 Hedging_{i,t} + \beta_5 Candor_{it} + \beta_6 log(1 + SIZE_{i,t}) + \beta_7 ROA_{i,t} + \beta_8 MOMENTUM_{i,t} + \beta_9 ESGscore_{i,t} + \gamma_{Industry} + \delta_{Year} + \varepsilon_{i,t}$$

This model directly tests Hypothesis 1 by evaluating whether the individual linguistic traits underlying greenwashing are empirically meaningful. A significant negative association for one or more components would confirm their relevance. Additionally, the model contributes to Hypothesis 2 by identifying the specific forms of linguistic manipulation or vagueness to which investors penalise most strongly.

3.8.4. Model (M4) – Interactions Between Linguistic Traits and ESG Context

Model 4 investigates how combinations of linguistic and contextual factors influence investor reactions by introducing two interaction terms. The first is the interaction between buzzwords and sentiment, to determine whether the use of positive emotional language combined with vague promotional terms is perceived as insincere by investors. The second interaction links specificity with the firm's ESG score, to assess whether concrete language enhances perceived authenticity, particularly for firms already viewed as strong ESG performers. This model takes the following form:

$$CAR_{i,t} = \beta_0 + \beta_1 (Buzzwords \times Sentiment)_{i,t} + \beta_2 (Specificity \times ESGscore)_{i,t} + \beta_3 Hedging_{i,t} + \beta_4 Candor_{it} + \beta_5 log(1 + SIZE_{i,t}) + \beta_6 ROA_{i,t} + \beta_7 MOMENTUM_{i,t} + \gamma_{Industry} + \delta_{Year} + \varepsilon_{i,t}$$

A negative and significant coefficient on *buzzword* × *sentiment* would suggest that emotionally charged, vague language weakens market trust. Conversely, a positive coefficient on *specificity* × *ESG score* would indicate that detailed communication boosts investor confidence, particularly when paired with strong ESG fundamentals.

This model adds a nuanced layer to Hypothesis 1 by showing that greenwashing is not just about individual linguistic cues, but also how these cues interact to form a coherent (or incoherent) communication. It also refines Hypothesis 2 by showing that investor responses to ESG disclosures depend not only on the content (e.g. high ESG performance), but also on how it is communicated (e.g. clarity versus hype).

These four regression models provide the foundation for the empirical analysis. In the next chapter, we present and interpret the results, assessing whether linguistic indicators of greenwashing influence investor reactions to ESG disclosures.

4. Results

4.1. Overview

This section presents the empirical findings derived from the regression models described above. We report the estimated effects of the Greenwashing Score across different specifications and event windows, comparing both rule-based (NLP) and GPT-4-based approaches. The analysis follows the three CAR windows (mentioned in Section 3.6) to capture short- and medium-term investor reactions. We begin with Model 1, which tests the overall effect of the greenwashing score on abnormal returns, before turning to models that explore interaction effects and the impact of individual linguistic components.

4.2. Model 1 – Market Reaction to AI-Based Greenwashing Scores

The null hypothesis (H_0) tested here is that greenwashing, as measured by the GWS, has no effect on stock market reactions. Rejection of H_0 in favour of the alternative hypothesis (H_1) would require a statistically significant and negative coefficient on the greenwashing variable, indicating a market penalty for perceived insincerity in ESG reporting. Full regression outputs, including coefficients and significance levels, are provided in Appendix 9.

4.2.1. CAR2

In the short-term event window, both greenwashing measures yield negative coefficients: $\beta = -0.00513$ ($p = 0.560$) for the GPT-based score, and $\beta = -0.00245$ ($p = 0.854$) for the rule-based NLP score. While the direction of the estimates aligns with our hypothesis which suggests that higher greenwashing is associated with lower short-term returns, neither result is statistically significant. As such, we fail to reject the null hypothesis: there is no evidence of immediate market penalties for greenwashing.

4.2.2. CAR5

Results remain consistent over this event-window. The GPT-based score is associated with a coefficient of $\beta = -0.00679$ ($p = 0.583$), while the rule-based score yields $\beta = -0.00051$ ($p = 0.978$). Both estimates are negative, which indicates a persistent link between greenwashing and lower returns. However, the lack of statistical significance again prevents us from rejecting the null hypothesis.

4.2.3. CAR60

In the 60-day event window, the results remain directionally stable but statistically weak. The GPT-based score yields a coefficient of $\beta = -0.00889$ ($p = 0.798$), while the NLP-based score produces a nearly identical result of $\beta = -0.00435$ ($p = 0.935$). These outcomes suggest that even over a longer period, there is no meaningful or statistically significant relationship between greenwashing intensity and stock market performance. Thus, we find no evidence of a delayed market reaction to perceived greenwashing.

4.2.4. Interpretations

Across all three event windows, both greenwashing scores generated negative but statistically insignificant coefficients. While this pattern aligns with Hypothesis 2 (see Section 3.1.) in terms of direction, we cannot reject the null hypothesis (H_0). This suggests that perceived greenwashing, as captured by aggregated scores, does not systematically influence stock market reactions. These results imply that, at least in aggregate form, the Greenwashing Score does not carry enough explanatory power to systematically predict market behaviour in response to ESG disclosures. These inconclusive results also raise questions about Hypothesis 1, particularly with regard to the effectiveness of aggregated greenwashing metrics in capturing investor perception of ESG credibility. It is possible that investors respond more to specific cues or to particular contexts, and that aggregation may dilute those signals. Overall, these findings suggest that the market does not consistently recognise or penalise greenwashing when expressed through a global linguistic score, possibly because the signal is too weak or subtle, or is obscured by other confounding factors.

4.3. Model 2 – Interaction with ESG Performance

Model 2 tests whether the effect of greenwashing on stock market reactions depends on the firm's actual ESG performance. Full regression tables for M2 are available in Appendix 10.

4.3.1. CAR2

Within the short-term event window, both the GPT- and rule-based models generate interaction terms that are statistically insignificant. The GWS_gpt × ESGscore interaction has a coefficient of 0.00118 ($p = 0.779$), and the greenwashing_score × ESGscore interaction has a coefficient of –0.00165 ($p = 0.744$). It suggests that investors do not react significantly differently to greenwashing depending on ESG performance in the immediate aftermath of the disclosure.

4.3.2. CAR5

Within the five-day period, the interaction effects remain weak and statistically insignificant. The GPT-based interaction term has a coefficient of 0.00281 ($p = 0.635$) and the NLP-based term has a coefficient of 0.0116 ($p = 0.1028$). While the latter is close to significance, the positive sign contradicts expectations of harsher market reactions for high-rated ESG firms. This weakens the support for a penalty effect tied to perceived ESG hypocrisy.

4.3.3. CAR60

Over this window, the findings remain consistent. The interaction terms again fail to achieve statistical significance: –0.00139 ($p = 0.933$) for GWS_gpt × ESGscore, and –0.00699 ($p = 0.727$) for greenwashing_score × ESGscore. Despite the negative signs, the lack of statistical strength indicates that greenwashing is not more harshly punished over time when it comes from firms with stronger ESG reputations.

4.3.4. Interpretations

Model 2 examined whether investors penalise greenwashing more harshly when it comes from firms with strong ESG ratings. Across all three CAR windows, the interaction terms between greenwashing and ESG performance were statistically insignificant, with inconsistent coefficients. The rule-based model showed a marginally positive interaction in CAR5, challenging the hypothesis. These findings suggest that the market does not systematically react more negatively to greenwashing when it contradicts otherwise credible

ESG credentials. It could be either due to a lack of sensitivity or because prior ESG reputation softens the blow. Therefore, Hypothesis 2 remains unsupported in this case.

4.4. Model 3 – Disaggregated Linguistic Cues and Market Reaction

Model 3 tests whether the stock market responds more distinctly to individual linguistic features rather than to aggregated greenwashing scores (see Results in Appendix 11).

4.4.1. CAR2

The results show no significant effects among the GPT-derived components. Although the effect of specificity (NLP model) is not statistically significant at conventional levels (e.g. 5%), it is significant at the 10% threshold ($p = 0.078$). This suggests a potential association which is worth noting. Interestingly, hedging emerges as a significant positive predictor ($p = 0.039$). While high hedging typically reflects a lack of firm commitment, the positive coefficient suggests that investors may interpret such cautious language not as greenwashing, but rather as a sign of strategic risk awareness or realistic goal-setting.

4.4.2. CAR5

In the CAR5 window, most coefficients remain statistically insignificant. However, hedging in the rule-based specification, remains marginally positively significant ($p = 0.0624$). Specificity also maintains a near-significant relationship ($p = 0.0682$). These positive coefficients suggest that both cautious language and precise wording may be perceived positively by investors over a few trading days, possibly reflecting a preference for transparency balanced with prudent communication. The GPT-based features, in contrast, do not yield meaningful results in this window either.

4.4.3. CAR60

At the 60-day horizon, disaggregated components still fail to produce strong or consistent signals in the GPT-based model. However, in the rule-based model, specificity becomes statistically significant at the 5% level ($p = 0.0447$), suggesting that investors value clear ESG disclosures. Unlike in shorter windows, hedging is no longer statistically significant in the CAR60 model ($p = 0.1132$), suggesting that investor sensitivity to vague language may diminish over time. Sentiment, buzzwords, and candor remain consistently non-significant, suggesting these traits are less influential in shaping longer-term market reactions.

4.4.4. Interpretations

Overall, Model 3 gives partial support to Hypothesis 1 (see Section 3.1), showing that AI-based textual analysis can detect meaningful variation in ESG communication styles. Specifically, the repeated significance of features like specificity and hedging in the rule-based models suggests that certain linguistic traits elicit measurable investor responses. Interestingly, while hedging is often viewed as a potential greenwashing tactic, its positive and significant association with CAR suggests that investors may not perceive cautious language as deceptive, instead, it may signal prudence or risk awareness. Conversely, sentiment, candor, and buzzword usage appear less effective in distinguishing credible from insincere reports. These findings support the relevance of targeted linguistic indicators for identifying greenwashing, while also reinforcing the idea that the market response to ESG disclosures is subtle and context-dependent.

4.5. Model 4 – Interaction Effects

Model 4 explores potential interaction effects between key greenwashing dimensions in ESG reports. Specifically, we test (i) whether the combination of positive tone and buzzword intensity amplifies market skepticism (i.e., *buzzwords* × *sentiment*) and (ii) whether the market rewards higher specificity only when firms simultaneously exhibit higher ESG scores (*specificity* × *ESGscore*). Regression outputs are reported in full in Appendix 12.

4.5.1. CAR2

The results reveal no significant interaction effects in either specification. The coefficient for *buzzwords* × *sentiment* is consistently negative but far from significance (e.g., $p = 0.731$ in the GPT model). Similarly, the *specificity* × *ESGscore* interaction remains non-significant across both models. Overall, the market appears unresponsive in the very short term to these combinations of tone and content.

4.5.2. CAR5

In the five-day window, interaction terms continue to show no statistically significant effect. The *buzzwords* × *sentiment* coefficient remains negative in all models but again statistically irrelevant ($p = 0.654$ and 0.772). However, the *specificity* × *ESGscore* term approaches significance in the rule-based model ($p = 0.065$), suggesting that specificity (and not the lack of thereof) might generate positive reactions when paired with strong ESG fundamentals.

Hedging also shows borderline significance in the NLP-based model ($p = 0.057$), consistent with findings in Model 3.

4.5.3. *CAR60*

The CAR60 results mirror the earlier patterns: no interaction term reaches significance in either the GPT or rule-based models. Interestingly, while *specificity* × *ESGscore* remains positive, its coefficient lacks statistical strength ($p \approx 0.328$–0.386). Other variables such as tone, buzzwords, specificity, candor, and hedging, do not demonstrate independent or joint effects with any economic significance.

4.5.4. Interpretations

Model 4 provides no empirical support for the hypothesis, which says that interactions between tone and superficiality, or between clarity and ESG substance would moderate investor reactions. Despite the theoretical plausibility of these effects, the data reveal no statistically reliable relationships between linguistic cues and ESG fundamentals. At best, some weak signals point toward possible relevance of specificity in contexts of high ESGscore, but these remain inconclusive. The lack of significant interaction effects may indicate that investors either do not pick up on subtle linguistic nuances, or do not consider them meaningful when forming their reactions.

5. Discussion

5.1. Key Contributions

Our research offers a fresh perspective on existing literature by examining the credibility of ESG disclosures through linguistic analysis and artificial intelligence into the assessment of greenwashing risk. While past studies have focused on ESG content, performance metrics, or market reactions to scandals (Delmas & Burbano, 2011; Lyon & Maxwell, 2011; Akyildirim et al., 2023; Xu et al., 2025), few have explored how the style and tone of ESG language influence investor perceptions in routine contexts.

Building on previous work in financial text analysis and impression management (Cho et al., 2015; Melloni et al., 2017), we contribute to existing literature by developing a Greenwashing Score based on five linguistic dimensions: buzzword density, vagueness, sentiment tone, hedging language and lack of candor. These indicators are rooted in communication theory and stakeholder trust literature, and reflect the growing concerns of stakeholders and investors regarding overly glossy or ambiguous ESG communications (Boiral et al., 2021; Hassan, 2024).

What distinguishes our approach is the integration of GPT-4 as an AI-based co-analyst, used to assess the same linguistic indicators through its own internal logic and interpretative reasoning. This comparison between rule-based scoring and LLM-driven judgment provides a rare look into the alignment and divergence between traditional computational methods and the contextual understanding enabled by generative AI. Our results show that while GPT-4 places slightly more emphasis on specificity and less on buzzword usage than our initial model, both approaches converge on the same core dimensions of symbolic disclosure.

This study contributes to the literature by introducing a novel, language-based approach to assess ESG credibility. Using AI-driven textual analysis, it offers a replicable method to detect potential greenwashing in sustainability reporting, independent of scandals or third-party ratings. By linking these linguistic cues to stock market reactions, we provide a new lens for understanding how communication style may influence investor perception, even if not systematically priced in.

Finally, this thesis adopts an interdisciplinary approach by combining methods and insights from finance, natural language processing, corporate governance and business ethics. Building on earlier insights from the impression management literature, it suggests that language is not merely a means of communication, but also a strategic tool for shaping stakeholder trust and legitimacy (Cho et al., 2015; Merkl-Davies & Brennan, 2008).

5.2. Limitations

Despite its contributions, this study has several limitations. First, the Greenwashing Score relies on textual features that, while quantifiable, remain interpretative. This creates a risk of false positives, whereby reports flagged as insincere by the model may actually reflect genuine, forward-looking communication. This highlights the need for prudence when interpreting linguistic indicators in isolation and the importance of supplementing text-based analysis with other data sources or contextual information.

Second, while GPT-4 adds valuable perspective, it presents challenges. Despite its impressive linguistic capabilities, it is a general-purpose model that has not been fine-tuned on ESG-specific corpora, and as such, its understanding of sustainability discourse may lack domain sensitivity. Its responses are also highly prompt-dependent, meaning that slight variations in phrasing can produce noticeably different outputs, which may affect replicability and consistency in analytical tasks. Additionally, the model operates as a black box, with no transparent logic behind its internal weighting or interpretive reasoning. This raises concerns about knowledge transparency, particularly in contexts requiring high levels of accountability, such as financial analysis or sustainability auditing. Its assessments may also reflect biases embedded in corporate discourse, potentially reinforcing symbolic language instead of detecting it.

Third, the event study design captures correlations but not causality. Despite the inclusion of control variables such as firm size, profitability, and ESG ratings to mitigate confounding effects, residual noise remains and it is difficult to isolate the precise effect of greenwashing-related language on investor behaviour.

Finally, the score's weighting scheme, although informed by theory and literature, remains partly subjective. While robustness checks were performed, other weighting structures might yield different results. Future research could use machine learning to refine the weighting and improve predictive accuracy.

5.3. Future Research Directions

This thesis opens several promising avenues for future research on AI-based greenwashing detection. One direction is the creation of ESG-specific language models, such as generative LLMs fine-tuned on ESG disclosures and regulatory texts. These could enhance contextual understanding, reduce hallucinations, and improve the reliability of ESG language analysis.

Another opportunity lies in integrating non-textual data, such as carbon footprints, executive pay, or satellite imagery, to assess the alignment between ESG narratives and actual impact. This multimodal approach could help to distinguish between firms that communicate in an aspirational way and those that are misleading.

Finally, it would be interesting to conduct long-term studies to examine whether persistently high greenwashing scores across several reporting periods act as an indicator of reputational damage, regulatory scrutiny or long-term financial underperformance. Such links would reinforce the strategic importance of sincere and transparent ESG communication from both an ethical and a risk management perspective.

Taken together, these future directions suggest that the intersection of AI, textual analysis, and sustainable finance is still in its early stages, but has considerable potential for innovation, accountability, and real-world impact.

6. Conclusion

In response to growing demand for credible ESG communication, this thesis set out to examine whether AI-driven linguistic analysis can enhance the detection of greenwashing in corporate disclosures (H1), and whether such detection influences investor behaviour (H2). To explore these questions, we developed a Greenwashing Score based on five textual indicators, combining rule-based NLP methods with GPT-4 scoring, and tested their financial relevance using event-study regressions across multiple CAR windows.

The results provide a nuanced but informative answer to our research question. Regarding our first hypothesis (H1), the constructed Greenwashing Score, particularly in its rule-based form, proved capable of capturing meaningful variation in ESG narrative styles. Specifically, features such as specificity and hedging consistently showed associations with market reactions, suggesting that AI-driven language analysis can indeed detect communicative traits relevant to perceptions of sincerity. In contrast, GPT-4's contextual scores were less effective, indicating that structured, theory-driven models may currently outperform generative AI in capturing market-relevant greenwashing cues. Overall, these findings provide only partial support for H2, suggesting that in the absence of reputational shocks or contextual triggers, linguistic cues alone are unlikely to systematically influence investor behaviour. Contrary to expectations, linguistic signals, such as cautious language (hedging), were not penalised, and in some cases even associated with slightly positive market responses. This suggests that investors do not systematically react negatively to perceived greenwashing in ESG disclosures unless accompanied by reputational shocks or broader contextual factors.

In conclusion, this study demonstrates that while AI-based detection of greenwashing is feasible and potentially relevant to markets, it remains an evolving tool. While not yet sufficient to predict investor reactions on its own, such an approach offers a scalable and transparent tool to assess narrative credibility. By laying the groundwork for more refined, linguistically informed ESG analysis, this thesis opens new avenues for research at the intersection of AI, sustainability data, and market behaviour, ultimately contributing to stronger accountability in sustainable finance.

7. Bibliography

Akyildirim, E., Corbet, S., Ongena, S., & Oxley, L. (2023). *Greenwashing: Do Investors, Markets and Boards Really Care?* (SSRN Scholarly Paper 4598145). Social Science Research Network. https://doi.org/10.2139/ssrn.4598145

Amer, S., & Ezz, M. (2023). The Impact of Greenwashing on Brand Reputation, Brand Credibility, and Green Brand Equity: Evidence from the Household Appliances Market. *International Journal of Marketing Studies*, *15*(2), Article 2. https://doi.org/10.5539/ijms.v15n2p84

Arouri, M., El Ghoul, S., & Gomes, M. (2021). Greenwashing and product market competition. *Finance Research Letters*, *42*, 101927. https://doi.org/10.1016/j.frl.2021.101927

Ballan, B., & Czarnezki, J. J. (2024). *Disclosure, Greenwashing, and the Future of ESG Litigation* (SSRN Scholarly Paper 4938648). https://doi.org/10.2139/ssrn.4938648

Banerjee, S., Agarwal, A., & Singla, S. (2024). *LLMs Will Always Hallucinate, and We Need to Live With This* (arXiv:2409.05746). arXiv. https://doi.org/10.48550/arXiv.2409.05746

Berg, F., Kölbel, J. F., & Rigobon, R. (2022). Aggregate Confusion: The Divergence of ESG Ratings*. *Review of Finance*, *26*(6), 1315–1344. https://doi.org/10.1093/rof/rfac033

Big shifts, small steps. (2022). KPMG. Retrieved March 19, 2025, from https://kpmg.com/xx/en/our-insights/esg/survey-of-sustainability-reporting-2022.html

Boiral, O. (2013). Sustainability reports as simulacra? A counter-account of A and A+ GRI reports. *Accounting, Auditing & Accountability Journal*, *26*(7), 1036–1071. https://doi.org/10.1108/AAAJ-04-2012-00998

Boiral, O., Talbot, D., Brotherton, M.-C., & Heras-Saizarbitoria, I. (2021). Sustainability rating and moral fictionalism: Opening the black box of nonfinancial agencies. *Accounting, Auditing & Accountability Journal*, *34*(8), 1740–1768. https://doi.org/10.1108/AAAJ-12-2019-4356

Bridge, A. (2024, March 7). Company level disclosure key to financing green transition, says ESMA chair. *Reuters*. https://www.reuters.com/sustainability/sustainable-finance-reporting/company-level-disclosur e-key-financing-green-transition-says-esma-chair-2024-03-07/

Chen, S., Song, Y., & Gao, P. (2023). Environmental, social, and governance (ESG) performance and financial outcomes: Analyzing the impact of ESG on financial performance. *Journal of Environmental Management*, *345*, 118829. https://doi.org/10.1016/j.jenvman.2023.118829

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. https://doi.org/10.1016/j.aos.2014.12.003

Christensen, L. T., Morsing, M., & Thyssen, O. (2013). CSR as Aspirational Talk. *Organization*, *20*(3), 372–393. https://doi.org/10.1177/1350508413478310

Chuang, M., Chuang, G., Chuang, C., & Chuang, J. (2025). *Judging It, Washing It: Scoring and Greenwashing Corporate Climate Disclosures using Large Language Models* (arXiv:2502.15094). arXiv. https://doi.org/10.48550/arXiv.2502.15094

Cifrino, D. (2023, June 29). The Rise of International ESG Disclosure Standards. *The Harvard Law School Forum on Corporate Governance*. https://corpgov.law.harvard.edu/2023/06/29/the-rise-of-international-esg-disclosure-standards/

Dame, M. C. W. | U. of N. (n.d.). *Loughran-McDonald Master Dictionary w/ Sentiment Word Lists*. Software Repository for Accounting and Finance. Retrieved May 4, 2025, from https://sraf.nd.edu/loughranmcdonald-master-dictionary/

de Freitas Netto, S. V., Sobral, M. F. F., Ribeiro, A. R. B., & Soares, G. R. da L. (2020). Concepts and forms of greenwashing: A systematic review. *Environmental Sciences Europe*, *32*(1), 19. https://doi.org/10.1186/s12302-020-0300-3

De Jong, M. D. T., Harkink, K. M., & Barth, S. (2018). Making Green Stuff? Effects of Corporate Greenwashing on Consumers. *Journal of Business and Technical Communication*, *32*(1), 77–112. https://doi.org/10.1177/1050651917729863

Delmas, M. A., & Burbano, V. C. (2011). The Drivers of Greenwashing. *California Management Review*, *54*(1), 64–87. https://doi.org/10.1525/cmr.2011.54.1.64

Deutsche Bank-owned asset manager DWS fined $27 million for greenwashing | Reuters. (2025). Retrieved May 20, 2025, from https://www.reuters.com/sustainability/german-asset-manager-dws-fined-25-mln-eur-greenwashing-case-2025-04-02/

Dhaliwal, D., Li, O. Z., Tsang, A., & Yang, Y. G. (2014). Corporate social responsibility disclosure and the cost of equity capital: The roles of stakeholder orientation and financial transparency. *Journal of Accounting and Public Policy*, *33*(4), 328–355. https://doi.org/10.1016/j.jaccpubpol.2014.04.006

Dong, M. M., Stratopoulos, T. C., & Wang, V. X. (2024). A scoping review of ChatGPT research in accounting and finance. *International Journal of Accounting Information Systems*, *55*, 100715. https://doi.org/10.1016/j.accinf.2024.100715

Dzieliński, M., Wagner, A. F., & Zeckhauser, R. J. (2017). *Straight Talkers and Vague Talkers: The Effects of Managerial Style in Earnings Conference Calls* (SSRN Scholarly Paper 2971822). Social Science Research Network. https://papers.ssrn.com/abstract=2971822

Ernst, C. A., Kunz, J., Fischer, T. M., & Ludwig, L. M. (2025). Investors' reactions to CSR reputation and disclosure assurance: An experimental analysis. *Journal of Management and Governance*. https://doi.org/10.1007/s10997-024-09734-8

Explainable AI for Greenwashing Detection | Ocean Business b2match 2025. (2025). Retrieved March 20, 2025, from https://www.b2match.com/e/ocean-business-2025/opportunities/UGFydGljaXBhdGlvbk9wcG9ydHVuaXR5OjE1MDg0Mg==

Flammer, C. (2013). Corporate Social Responsibility and Shareholder Reaction: The Environmental Awareness of Investors. *Academy of Management Journal*, *56*(3), 758–781. https://doi.org/10.5465/amj.2011.0744

Gehrig, N., & Moreno, A. (2023). *An Exploration of Greenwashing Risks in Investment Fund Disclosures: An Investor Perspective*. CFA Institute. https://doi.org/10.56227/23.1.19

Global Sustainable Investment Review finds US$30 trillion invested in sustainable assets. (2023). Retrieved February 12, 2025, from https://www.gsi-alliance.org/global-sustainable-investment-review-finds-us30-trillion-invested-in-sustainable-assets/

Green claims—European Commission. (n.d.). Retrieved January 14, 2025, from https://environment.ec.europa.eu/topics/circular-economy/green-claims_en

Green Guides | Federal Trade Commission. (n.d.). Retrieved January 14, 2025, from https://www.ftc.gov/news-events/topics/truth-advertising/green-guides

GreenWatch Methodology—GreenWatch. (2021, August 13). https://greenwatch.ai/greenwatch-methodology/

Habib, A., Oláh, J., Khan, M. H., & Luboš, S. (2025). Does integration of ESG disclosure and green financing improve firm performance: Practical applications of stakeholders theory. *Heliyon*, *11*(4), e41996. https://doi.org/10.1016/j.heliyon.2025.e41996

Hahn, R., & Lülfs, R. (2014). Legitimizing Negative Aspects in GRI-Oriented Sustainability Reporting: A Qualitative Analysis of Corporate Disclosure Strategies. *Journal of Business Ethics*, *123*(3), 401–420. https://ideas.repec.org//a/kap/jbuset/v123y2014i3p401-420.html

Hassan, S. M. (2024). Greenwashing in ESG: Identifying and Addressing False Claims of Sustainability. *Journal of Business and Strategic Management*, *9*(8), 90–105. https://ideas.repec.org//a/bhx/ojjbsm/v9y2024i8p90-105id2390.html

Hu, P., Li, X., Li, N., Wang, Y., & Wang, D. D. (2024). Peeking into Corporate Greenwashing through the Readability of ESG Disclosures. *Sustainability*, *16*(6), Article 6. https://doi.org/10.3390/su16062571

Huang, F., Kwak, H., Park, K., & An, J. (2024). *ChatGPT Rates Natural Language Explanation Quality Like Humans: But on Which Scales?* (arXiv:2403.17368). arXiv. https://doi.org/10.48550/arXiv.2403.17368

Huang, J., Li, Y., & Han, F. (2024). Walk well and talk well: Impact of the consistency of ESG performance and disclosure on firms' stock price crash risk. *International Review of Economics & Finance*, *93*, 1154–1174. https://doi.org/10.1016/j.iref.2024.03.067

Kang, H., & Kim, J. (2022). Analyzing and Visualizing Text Information in Corporate Sustainability Reports Using Natural Language Processing Methods. *Applied Sciences*, *12*(11), Article 11. https://doi.org/10.3390/app12115614

Kim, J., Fairclough, S., & Dibrell, C. (2017). Attention, Action, and Greenwash in Family-Influenced

Firms? Evidence From Polluting Industries. *Organization & Environment*, *30*(4), 304–323. https://doi.org/10.1177/1086026616673410

Kotte, A., Woodcock, D., & Guynn, J. (2019, August 12). Managing Legal Risks from ESG Disclosures. *The Harvard Law School Forum on Corporate Governance*. https://corpgov.law.harvard.edu/2019/08/12/managing-legal-risks-from-esg-disclosures/

Krüger, P. (2015). Corporate goodness and shareholder wealth. *Journal of Financial Economics*, *115*(2), 304–329. https://doi.org/10.1016/j.jfineco.2014.09.008

Leippold, M., Vaghefi, S. A., Stammbach, D., Muccione, V., Bingler, J., Ni, J., Colesanti-Senni, C., Wekhof, T., Schimanski, T., Gostlow, G., Yu, T., Luterbacher, J., & Huggel, C. (2024). *Automated Fact-Checking of Climate Change Claims with Large Language Models* (arXiv:2401.12566). arXiv. https://doi.org/10.48550/arXiv.2401.12566

Li, R., Wu, H., Zhong, M., & Lan, H. (2024). Strategic tone management in ESG reports and ESG risk. *International Review of Financial Analysis*, *96A*, 103618–103618. https://doi.org/10.1016/j.irfa.2024.103618

Li, T., Shu, X., & Liao, G. (2024). Does corporate greenwashing affect investors' decisions? *Finance Research Letters*, *67*, 105877. https://doi.org/10.1016/j.frl.2024.105877

Lin, W. L., Chong, S. C., Pek, C. K., Yong, J. Y., Ming, K. L. Y., & XeChung, N. L. (2025). *The Impact of Greenwashing: Risks and Implications for Corporate Performance and Stakeholder Trust*. 77–94. https://doi.org/10.2991/978-2-38476-358-0_8

Loughran, T., & Mcdonald, B. (2011). When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *The Journal of Finance*, *66*(1), 35–65. https://doi.org/10.1111/j.1540-6261.2010.01625.x

Lyon, T. P., & Maxwell, J. W. (2011). Greenwash: Corporate Environmental Disclosure under Threat of Audit. *Journal of Economics & Management Strategy*, *20*(1), 3–41. https://doi.org/10.1111/j.1530-9134.2010.00282.x

Melloni, G., Caglio, A., & Perego, P. (2017). Saying more with less? Disclosure conciseness, completeness and balance in Integrated Reports. *Journal of Accounting and Public Policy*, *36*(3), 220–238. https://doi.org/10.1016/j.jaccpubpol.2017.03.001

Melloni, G., Stacchezzini, R., & Lai, A. (2016). The tone of business model disclosure: An impression management analysis of the integrated reports. *Journal of Management & Governance*, *20*(2), 295–320. https://doi.org/10.1007/s10997-015-9319-z

Merkl-Davies, D. M., & Brennan, N. M. (2008). *Discretionary Disclosure Strategies in Corporate Narratives: Incremental Information or Impression Management?* (SSRN Scholarly Paper 1089447). Social Science Research Network. https://papers.ssrn.com/abstract=1089447

Merritt, A. C., Effron, D. A., & Monin, B. (2010). Moral Self-Licensing: When Being Good Frees Us to Be Bad. *Social and Personality Psychology Compass*, *4*(5), 344–357. https://doi.org/10.1111/j.1751-9004.2010.00263.x

Müller, S. (2024, September 5). *Understanding Cumulative Abnormal Return (CAR) in Finance*. Event Study. https://eventstudy.de/blog/cumulative-abnormal-return/

Nicolas, M. L. D., Desroziers, A., Caccioli, F., & Aste, T. (2023). *ESG Reputation Risk Matters: An Event Study Based on Social Media Data* (arXiv:2307.11571). arXiv. https://doi.org/10.48550/arXiv.2307.11571

Ong, K., Mao, R., Satapathy, R., Filho, R. S., Cambria, E., Sulaeman, J., & Mengaldo, G. (2025). Explainable natural language processing for corporate sustainability analysis. *Information Fusion*, *115*, 102726. https://doi.org/10.1016/j.inffus.2024.102726

OpenAI. (2023). *ChatGPT (GPT-4)* [Large language model]. https://chat.openai.com/

Parguel, B., Benoît-Moreau, F., & Larceneux, F. (2011). How Sustainability Ratings Might Deter 'Greenwashing': A Closer Look at Ethical Corporate Communication. *Journal of Business Ethics*, *102*(1), 15–28. https://doi.org/10.1007/s10551-011-0901-2

PricewaterhouseCoopers. (2024.). *94% of investors believe corporate reporting on sustainability performance contains unsupported claims: PwC 2023 Global Investor Survey*. PwC. Retrieved March 7, 2025, from https://www.pwc.com/th/en/press-room/press-release/2024/press-release-10-01-24-en.html

Quang Trinh, V., Duong Cao, N., Li, T., & Elnahass, M. (2023). Social capital, trust, and bank tail risk: The value of ESG rating and the effects of crisis shocks. *Journal of International Financial Markets, Institutions and Money*, *83*, 101740.

https://doi.org/10.1016/j.intfin.2023.101740

Rouen, E., Sachdeva, K., & Yoon, A. (2022). *The Evolution of ESG Reports and the Role of Voluntary Standards—Working Paper—Faculty & Research—Harvard Business School*. https://www.hbs.edu/faculty/Pages/item.aspx?num=63173

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*, 239–252. https://doi.org/10.1002/bse.1912

Stuart, A. C., Bedard, J. C., & Clark, C. E. (2021). Corporate Social Responsibility Disclosures and Investor Judgments in Difficult Times: The Role of Ethical Culture and Assurance. *Journal of Business Ethics*, *171*(3), 565–582. https://ideas.repec.org//a/kap/jbuset/v171y2021i3d10.1007_s10551-020-04454-z.html

Surana, S., Dembla, S., & Bihani, P. (2022). Identifying Contradictions in the Legal Proceedings Using Natural Language Models. *SN Computer Science*, *3*(3), 187. https://doi.org/10.1007/s42979-022-01075-3

Tang, K.-Y., Chang, C.-Y., & Hwang, G.-J. (2023). Trends in artificial intelligence-supported e-learning: A systematic review and co-citation network analysis (1998–2019). *Interactive Learning Environments*, *31*(4), 2134–2152. https://doi.org/10.1080/10494820.2021.1875001

Team, T. C. (2025, April 3). How AI is exposing greenwashing and false ESG claims. *The Choice by ESCP*. https://thechoice.escp.eu/tomorrow-choices/greenwashing-in-corporate-sustainability-how-ai-is-exposing-false-esg-claims/

Terrachoice. (2009). *The Seven Sins of Greenwashing | Sustainability | Business*. Scribd. https://www.scribd.com/document/14475597/The-Seven-Sins-of-Greenwashing

Theil, K., & Stuckenschmidt, H. (2020). Predicting Modality in Financial Dialogue. In D. M. El-Haj, D. V. Athanasakou, D. S. Ferradans, D. C. Salzedo, D. A. Elhag, D. H. Bouamor, D. M. Litvak, D. P. Rayson, D. G. Giannakopoulos, & N. Pittaras (Eds.), *Proceedings of the 1st Joint Workshop on Financial Narrative Processing and MultiLing Financial Summarisation* (pp. 226–234). COLING. https://aclanthology.org/2020.fnp-1.35/

Toffel, M., & Marquis, C. (2012). *When Do Firms Greenwash? Corporate Visibility, Civil Society Scrutiny, and Environmental Disclosure*. https://www.academia.edu/2857238/When_Do_Firms_Greenwash_Corporate_Visibility_Civil_Society_Scrutiny_and_Environmental_Disclosure

Using Artificial Intelligence Used to Find Corporate Greenwashers: Green Insight—Bloomberg. (2021). Retrieved November 4, 2024, from https://www.bloomberg.com/news/articles/2021-08-11/using-artificial-intelligence-used-to-find-corporate-greenwashers-green-insight?embedded-checkout=true

Vereb, M. (2024). *The Dangers of Greenwashing: How Companies Mislead Consumers*. https://www.arbor.eco/blog/the-greenwashers

Vinella, A., Capetz, M., Pattichis, R., Chance, C., Ghosh, R., & Chang, K.-W. (2024). *Leveraging Language Models to Detect Greenwashing* (arXiv:2311.01469). arXiv. https://doi.org/10.48550/arXiv.2311.01469

VW "dieselgate" fraud: Timeline of a scandal. (2019). Retrieved November 21, 2025, from https://techxplore.com/news/2019-09-vw-dieselgate-fraud-timeline-scandal.html

Walker, K., & Wan, F. (2012). The Harm of Symbolic Actions and Green-Washing: Corporate Actions and Communications on Environmental Performance and Their Financial Implications. *Journal of Business Ethics*, *109*(2), 227–242. https://doi.org/10.1007/s10551-011-1122-4

Wei, J., Tay, Y., Bommasani, R., Raffel, C., Zoph, B., Borgeaud, S., Yogatama, D., Bosma, M., Zhou, D., Metzler, D., Chi, E. H., Hashimoto, T., Vinyals, O., Liang, P., Dean, J., & Fedus, W. (2022). *Emergent Abilities of Large Language Models* (arXiv:2206.07682). arXiv. https://doi.org/10.48550/arXiv.2206.07682

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q., & Zhou, D. (2023). *Chain-of-Thought Prompting Elicits Reasoning in Large Language Models* (arXiv:2201.11903). arXiv. https://doi.org/10.48550/arXiv.2201.11903

Xu, C., Miao, Y., Xiao, Y., & Lin, C. (2025). *DeepGreen: Effective LLM-Driven Green-washing Monitoring System Designed for Empirical Testing -- Evidence from China* (arXiv:2504.07733). arXiv. https://doi.org/10.48550/arXiv.2504.07733

Xu, M., Tse, Y. K., Geng, R., Liu, Z., & Potter, A. (2025). Greenwashing and market value of firms: An empirical study. *International Journal of Production Economics*, *284*, 109606. https://doi.org/10.1016/j.ijpe.2025.109606

Yang, J. (2019). Analysis of Hedges in English Advertising from the Perspective of Cooperative Principle. *Open Access Library Journal*, *6*(6), Article 6. https://doi.org/10.4236/oalib.1105441

Zhang, D. (2024). The pathway to curb greenwashing in sustainable growth: The role of artificial intelligence. *Energy Economics*, *133*, 107562. https://doi.org/10.1016/j.eneco.2024.107562

Zou, Y., Shi, M., Chen, Z., Deng, Z., Lei, Z., Zeng, Z., Yang, S., Tong, H., Xiao, L., & Zhou, W. (2023). *ESGReveal: An LLM-based approach for extracting structured data from ESG reports* (arXiv:2312.17264). arXiv. https://doi.org/10.48550/arXiv.2312.17264

8. Appendices

- **Appendix 1 - Greenwashing Score Component's Computation (Section 3.3)**

Component	Description	Computation Method	Expected Greenwashing Pattern
Sentiment (Tone)	Measures overall tone (positive/negative) of ESG disclosures	$\text{Sentiment} = \frac{(Positive\ Words - Negative\ Words)}{Total\ Words}$	Overly positive tone may suggest greenwashing
Specificity	Degree of concreteness vs. vagueness in claims	$\text{Specificity} = \frac{Numerical\ values}{(Vague\ terms + 1)}$	Lower specificity suggests vague, non-committal ESG claims
Buzzword Density (per 1000 words)	Proportion of ESG-related buzzwords vs. total words	$\text{Buzzword density} = \frac{Buzzwords}{Total\ Words} \times 1000$	High density can be a sign of superficial commitment
Hedging Language	Degree of uncertainty or cautious phrasing	$\text{Hedging density} = \frac{Hedging\ terms}{Total\ Words} \times 1000$	High hedging indicates lack of firm commitment
Candor	Openness and balance (mention of challenges and negative outcomes)	$\text{Candor} = \frac{Candor{-}related\ words}{Total\ Words} \times 1000$	Lower candor may indicate image management

- **Appendix 2 - Correlation matrix (Section 3.4.4.1)**

	buzzwords_score_industry_norm	Specificity_Score	sentiment_risk_score	hedging_density_norma	candor_risk_component
buzzwords_score_industry_norm	1.00000000	-0.08101054	0.17342425	0.03153906	0.03452062
Specificity_Score	-0.08101054	1.00000000	0.28162253	0.44969397	-0.03514049
sentiment_risk_score	0.17342425	0.28162253	1.00000000	0.14854353	0.00821424
hedging_density_norma	0.03153906	0.44969397	0.14854353	1.00000000	-0.07221105
candor_risk_component	0.03452062	-0.03514049	0.00821424	-0.07221105	1.00000000

- **Appendix 3 - Contribution of Each Component to the GWS (Section 3.4.4.2)**

	buzzwords_score_industry_norm	Specificity_Score	sentiment_risk_score	hedging_density_norma	candor_risk_component	greenwashing_score
buzzwords_score_industry_norm	1.00000000	-0.08101054	0.17342425	0.03153906	0.03452062	0.8444052
Specificity_Score	-0.08101054	1.00000000	0.28162253	0.44969397	-0.03514049	0.3675037
sentiment_risk_score	0.17342425	0.28162253	1.00000000	0.14854353	0.00821424	0.4581213
hedging_density_norma	0.03153906	0.44969397	0.14854353	1.00000000	-0.07221105	0.3656178
candor_risk_component	0.03452062	-0.03514049	0.00821424	-0.07221105	1.00000000	0.1114058
greenwashing_score	0.84440518	0.36750366	0.45812134	0.36561777	0.11140579	1.0000000

- **Appendix 4 - Multicollinearity (Section 3.4.4.3)**

```
> model <- lm(greenwashing_score ~ buzzwords_score_industry_norm + Specificity_Score + sentiment_risk_score + hedging_density_norma + candor_risk_component, data = gws)
> vif(model) # VIF < 5 is generally acceptable
buzzwords_score_industry_norm             Specificity_Score          sentiment_risk_score         hedging_density_norma         candor_risk_component
                     1.047828                      1.362890                      1.129073                      1.258010                      1.006164
```

- **Appendix 5 - Prompt used to obtainChat-GPT's own greenwashing score (Section 3.5)**

Context:

You are a sustainability analyst. Your task is to assess the level of greenwashing in a company's ESG report. The full report is provided below.

Carefully evaluate the report based on five key dimensions of potential greenwashing. For each dimension:

- Assign a score between 0 and 1, where:
  0 = no signs of greenwashing
  1 = strong signs of greenwashing
- Briefly explain the reasoning behind your score.
  Please think step by step, as a human analyst would do.

Important:

No specific weights are given for the five dimensions. You are free to judge how much weight to give each one when forming your overall assessment. The goal is to reflect how a thoughtful analyst would intuitively assess the report's sincerity.

Five Dimensions to Evaluate:

1. Buzzword Usage
   Does the report rely heavily on trendy sustainability terms (e.g., "green," "eco-friendly," "net zero") without clear explanation or backing?
2. Specificity (or Vagueness)
   Are the claims supported by measurable data, timelines, or concrete goals? Or are they vague, aspirational, or generic?
3. Sentiment Tone
   Is the tone overly positive, promotional, or unbalanced? Does it sound like marketing rather than honest reporting?
4. Hedging Language
   Does the report use cautious, indirect, or non-committal wording (e.g., "may," "hope to," "intend to") that reduces accountability?
5. Candor (Transparency about Shortcomings)
   Does the report openly acknowledge ESG challenges, missed goals, or areas for improvement? Or does it present only successes?

Final Greenwashing Risk Score (0–1):

Based on your assessment of the five dimensions, assign an overall greenwashing score between 0 and 1. Summarize your reasoning across dimensions before giving the final score. Use your own judgment to determine how much each category should influence your overall conclusion.

- **Appendix 6 - Chat-GPT's weighting scheme (Section 3.5.3)**

Component	Transformed Metric	Weight	Rationale (based on GPT-4 behavior and explanations)
Specificity	`1 - specificity_score`	0.30	GPT-4 consistently emphasized the importance of measurable data
Tone (Sentiment)	`tone_score`	0.20	Overly optimistic tone was often flagged as problematic
Hedging	`hedging_score`	0.20	Conditional phrasing triggered moderate concern
Buzzwords	`buzzword_score`	0.15	Buzzwords were noted, but not always penalized if context was rich
Candor	`1 - candor_score`	0.15	Lack of transparency was penalized, but GPT-4 sometimes gave benefit of the doubt

- **Appendix 7 - Pearson Correlation (Section 3.5.4)**

```
> # Pearson correlation
> cor_value_ <- cor(merged_df$greenwashing_score, merged_df$GWS_gpt, use = "complete.obs", method = "pearson")
> round(cor_value_, 3)
[1] 0.276
```

- **Appendix 8 - Component-level comparisons (Section 3.5.4)**

```
> summary(Comparison_df[, grep("_Diff$", names(Comparison_df))])
 Buzzword_Diff      Specificity_Diff   Sentiment_Diff     Hedging_Diff      Candor_Diff         GWS_Diff
 Min.   :-1.11200   Min.   :-0.45548   Min.   :-0.3207   Min.   :-0.5500   Min.   :-1.0000   Min.   :-0.23391
 1st Qu.:-0.28304   1st Qu.: 0.02307   1st Qu.: 0.1905   1st Qu.: 0.4921   1st Qu.:-0.1878   1st Qu.: 0.01621
 Median :-0.19548   Median : 0.15063   Median : 0.2250   Median : 0.5316   Median :-0.1110   Median : 0.06900
 Mean   :-0.18476   Mean   : 0.11612   Mean   : 0.2269   Mean   : 0.5182   Mean   :-0.1321   Mean   : 0.07659
 3rd Qu.:-0.07928   3rd Qu.: 0.24031   3rd Qu.: 0.2617   3rd Qu.: 0.5697   3rd Qu.:-0.0570   3rd Qu.: 0.13090
 Max.   : 1.00000   Max.   : 1.00000   Max.   : 1.0000   Max.   : 0.8148   Max.   : 0.9423   Max.   : 0.76739
                    NA's   :7          NA's   :2                                              NA's   :9
```

- **Appendix 8 bis - Pearson Correlation Matrix Between NLP and GPT Scores (Exploratory)**

	Buzzword_NLP	Specificity_NLP	Sentiment_NLP	Hedging_NLP	Candor_NLP	GWS_NLP	Buzzword_GPT	Specificity_GPT	Sentiment_GPT	Hedging_GPT	Candor_GPT	GWS_GPT
Buzzword_NLP	1.00000000	0.081661650	0.16290204	-0.01971441	-0.038637995	0.84440518	0.32068412	0.116503158	-0.01679967	-0.051635056	0.081466423	-0.05591237
Specificity_NLP	0.08166165	1.000000000	-0.28162253	-0.44956944	-0.035763559	-0.36750366	0.03934162	0.614045064	-0.13441785	-0.033160548	-0.005379288	-0.63315611
Sentiment_NLP	0.16290204	-0.281622535	1.00000000	0.12553828	-0.010732961	0.45812134	0.07246597	-0.168532821	0.26801795	-0.081030761	-0.098769586	0.23619827
Hedging_NLP	-0.01971441	-0.449569442	0.12553828	1.00000000	0.067134626	0.36561777	0.15901570	-0.199977725	0.04172747	0.170930884	0.065335556	0.42541628
Candor_NLP	-0.03863800	-0.035763559	-0.01073296	0.06713463	1.000000000	-0.11140579	0.01147742	-0.002631895	-0.02976873	0.009216738	0.042385054	0.01393159
GWS_NLP	0.84440518	-0.367503664	0.45812134	0.36561777	-0.111405793	1.00000000	0.30885677	-0.153197078	0.09311781	-0.011303002	0.059354029	0.27635781
Buzzword_GPT	0.32068412	0.039341621	0.07246597	0.15901570	0.011477416	0.30885677	1.00000000	0.272772324	0.19555411	0.377568696	0.310779511	0.32786692
Specificity_GPT	0.11650316	0.614045064	-0.16853282	-0.19997772	-0.002631895	-0.15319708	0.27277232	1.000000000	-0.06911747	0.006158138	0.016652569	-0.70796561
Sentiment_GPT	-0.01679967	-0.134417845	0.26801795	0.04172747	-0.029768725	0.09311781	0.19555411	-0.069117468	1.00000000	0.216907903	0.093254419	0.26061377
Hedging_GPT	-0.05163506	-0.033160548	-0.08103076	0.17093088	0.009216738	-0.01130300	0.37756870	0.006158138	0.21690790	1.000000000	0.376406198	0.17685120
Candor_GPT	0.08146642	-0.005379288	-0.09876959	0.06533556	0.042385054	0.05935403	0.31077951	0.016652569	0.09325442	0.376406198	1.000000000	-0.09907660
GWS_GPT	-0.05591237	-0.633156113	0.23619827	0.42541628	0.013931589	0.27635781	0.32786692	-0.707965608	0.26061377	0.176851203	-0.099076602	1.00000000

- **Appendix 9 - Model 1 Regressions (Section 4.2)**
  - CAR2 (GPT vs. NLP)

```
Call:
lm(formula = CAR2 ~ GWS_gpt + log(1 + SIZE) + ROA + MOMENTUM +
    ESGscore + as.factor(Industry) + as.factor(year), data = merged_df)

Residuals:
      Min        1Q    Median        3Q       Max
-0.184277 -0.012923 -0.000193  0.012350  0.213333

Coefficients:
                                                Estimate Std. Error t value Pr(>|t|)
(Intercept)                                   -7.708e-03  9.522e-03  -0.809  0.41837
GWS_gpt                                       -5.133e-03  8.805e-03  -0.583  0.56003
log(1 + SIZE)                                 -4.054e-04  5.932e-04  -0.683  0.49449
ROA                                            5.382e-05  1.253e-04   0.430  0.66753
MOMENTUM                                      -3.314e-03  1.500e-03  -2.209  0.02730 *
ESGscore                                      -3.294e-04  3.404e-04  -0.968  0.33332
as.factor(Industry)Computers                   4.589e-03  4.038e-03   1.136  0.25601
as.factor(Industry)Durable manufacturers       8.205e-03  3.848e-03   2.132  0.03312 *
as.factor(Industry)Extractive industries       6.197e-03  5.239e-03   1.183  0.23704
as.factor(Industry)Financial institutions     -1.040e-03  4.446e-03  -0.234  0.81512
as.factor(Industry)Food                       -1.927e-03  4.522e-03  -0.426  0.67000
as.factor(Industry)Insurance and real estate   9.867e-06  5.424e-03   0.002  0.99855
as.factor(Industry)Mining and Construction    -7.104e-03  6.843e-03  -1.038  0.29936
as.factor(Industry)Others                      2.860e-03  2.116e-02   0.135  0.89247
as.factor(Industry)Pharmaceuticals             5.372e-03  4.736e-03   1.134  0.25684
as.factor(Industry)Retail                      7.221e-03  4.138e-03   1.745  0.08113 .
as.factor(Industry)Services                    6.810e-04  4.870e-03   0.140  0.88882
as.factor(Industry)Textiles, printing and publishing 3.112e-03 5.060e-03 0.615 0.53866
as.factor(Industry)Transportation              2.575e-03  4.673e-03   0.551  0.58168
as.factor(Industry)Utilities                   2.614e-03  4.907e-03   0.533  0.59437
as.factor(year)2008                            2.089e-02  7.389e-03   2.828  0.00475 **
as.factor(year)2009                            1.645e-02  7.260e-03   2.266  0.02357 *
as.factor(year)2010                            1.266e-02  6.788e-03   1.865  0.06240 .
as.factor(year)2011                            1.754e-02  6.774e-03   2.590  0.00969 **
as.factor(year)2012                            1.008e-02  6.510e-03   1.548  0.12173
as.factor(year)2013                            1.417e-02  6.431e-03   2.203  0.02771 *
as.factor(year)2014                            1.258e-02  6.375e-03   1.973  0.04870 *
as.factor(year)2015                            1.513e-02  6.343e-03   2.384  0.01722 *
as.factor(year)2016                            1.284e-02  6.385e-03   2.011  0.04447 *
as.factor(year)2017                            1.076e-02  6.301e-03   1.707  0.08798 .
as.factor(year)2018                            1.285e-02  6.231e-03   2.063  0.03927 *
as.factor(year)2019                            1.023e-02  6.175e-03   1.657  0.09767 .
as.factor(year)2020                            1.284e-02  6.997e-03   1.836  0.06660 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.02932 on 1634 degrees of freedom
Multiple R-squared:  0.03056,   Adjusted R-squared:  0.01157
F-statistic: 1.609 on 32 and 1634 DF,  p-value: 0.01714
```

```
Call:
lm(formula = CAR2 ~ greenwashing_score + log(1 + SIZE) + ROA +
    MOMENTUM + ESGscore + as.factor(Industry) + as.factor(year),
    data = merged_df)

Residuals:
      Min        1Q    Median        3Q       Max
-0.184389 -0.012746  0.000097  0.012496  0.212553

Coefficients:
                                                Estimate Std. Error t value Pr(>|t|)
(Intercept)                                   -1.150e-02  1.014e-02  -1.135  0.25654
greenwashing_score                             2.449e-03  1.333e-02   0.184  0.85429
log(1 + SIZE)                                 -3.247e-04  5.954e-04  -0.545  0.58555
ROA                                            4.968e-05  1.250e-04   0.398  0.69103
MOMENTUM                                      -3.290e-03  1.498e-03  -2.196  0.02821 *
ESGscore                                      -2.773e-04  3.384e-04  -0.819  0.41278
as.factor(Industry)Computers                   4.078e-03  4.108e-03   0.993  0.32109
as.factor(Industry)Durable manufacturers       8.008e-03  3.852e-03   2.079  0.03779 *
as.factor(Industry)Extractive industries       5.985e-03  5.217e-03   1.147  0.25146
as.factor(Industry)Financial institutions     -1.449e-03  4.429e-03  -0.327  0.74360
as.factor(Industry)Food                       -2.313e-03  4.654e-03  -0.497  0.61927
as.factor(Industry)Insurance and real estate  -2.935e-04  5.545e-03  -0.053  0.95780
as.factor(Industry)Mining and Construction    -7.351e-03  6.835e-03  -1.076  0.28230
as.factor(Industry)Pharmaceuticals             5.269e-03  4.718e-03   1.117  0.26429
as.factor(Industry)Retail                      7.210e-03  4.167e-03   1.730  0.08380 .
as.factor(Industry)Services                    8.417e-04  4.898e-03   0.172  0.86357
as.factor(Industry)Textiles, printing and publishing 2.851e-03 5.063e-03 0.563 0.57340
as.factor(Industry)Transportation              2.408e-03  4.653e-03   0.517  0.60493
as.factor(Industry)Utilities                   4.583e-04  5.059e-03   0.091  0.92784
as.factor(year)2008                            2.084e-02  7.361e-03   2.832  0.00469 **
as.factor(year)2009                            1.635e-02  7.226e-03   2.263  0.02377 *
as.factor(year)2010                            1.261e-02  6.757e-03   1.866  0.06221 .
as.factor(year)2011                            1.751e-02  6.743e-03   2.596  0.00952 **
as.factor(year)2012                            1.004e-02  6.480e-03   1.549  0.12151
as.factor(year)2013                            1.403e-02  6.404e-03   2.191  0.02862 *
as.factor(year)2014                            1.148e-02  6.355e-03   1.807  0.07091 .
as.factor(year)2015                            1.499e-02  6.325e-03   2.370  0.01791 *
as.factor(year)2016                            1.271e-02  6.357e-03   1.999  0.04582 *
as.factor(year)2017                            1.063e-02  6.277e-03   1.693  0.09061 .
as.factor(year)2018                            1.268e-02  6.211e-03   2.042  0.04128 *
as.factor(year)2019                            1.006e-02  6.159e-03   1.633  0.10269
as.factor(year)2020                            1.247e-02  6.965e-03   1.791  0.07354 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.02918 on 1626 degrees of freedom
  (9 observations deleted due to missingness)
Multiple R-squared:  0.03103,   Adjusted R-squared:  0.01256
F-statistic:  1.68 on 31 and 1626 DF,  p-value: 0.01122
```

  - CAR5 (GPT vs. NLP)

```
Call:
lm(formula = CAR5 ~ GWS_gpt + log(1 + SIZE) + ROA + MOMENTUM +
    ESGscore + as.factor(Industry) + as.factor(year), data = merged_df)

Residuals:
      Min        1Q    Median        3Q       Max
-0.228470 -0.020506  0.000651  0.020144  0.226270

Coefficients:
                                                Estimate Std. Error t value Pr(>|t|)
(Intercept)                                   -1.695e-02  1.336e-02  -1.269   0.2047
GWS_gpt                                       -6.785e-03  1.235e-02  -0.549   0.5828
log(1 + SIZE)                                  5.764e-04  8.321e-04   0.693   0.4886
ROA                                           -2.029e-04  1.757e-04  -1.154   0.2485
MOMENTUM                                      -3.141e-03  2.104e-03  -1.492   0.1358
ESGscore                                       4.664e-05  4.775e-04   0.098   0.9222
as.factor(Industry)Computers                  -4.171e-03  5.664e-03  -0.736   0.4616
as.factor(Industry)Durable manufacturers       2.912e-03  5.397e-03   0.540   0.5896
as.factor(Industry)Extractive industries       4.492e-03  7.349e-03   0.611   0.5412
as.factor(Industry)Financial institutions     -1.268e-02  6.237e-03  -2.033   0.0422 *
as.factor(Industry)Food                       -7.226e-03  6.343e-03  -1.139   0.2547
as.factor(Industry)Insurance and real estate  -2.960e-05  7.608e-03  -0.004   0.9969
as.factor(Industry)Mining and Construction    -1.066e-02  9.598e-03  -1.111   0.2668
as.factor(Industry)Others                     -3.436e-02  2.968e-02  -1.158   0.2471
as.factor(Industry)Pharmaceuticals             9.123e-04  6.644e-03   0.137   0.8908
as.factor(Industry)Retail                     -7.564e-05  5.804e-03  -0.013   0.9896
as.factor(Industry)Services                    8.547e-04  6.832e-03   0.125   0.9005
as.factor(Industry)Textiles, printing and publishing 3.614e-04 7.098e-03 0.051 0.9594
as.factor(Industry)Transportation             -7.401e-04  6.555e-03  -0.113   0.9101
as.factor(Industry)Utilities                  -5.758e-03  6.883e-03  -0.836   0.4030
as.factor(year)2008                            2.592e-02  1.036e-02   2.501   0.0125 *
as.factor(year)2009                            1.894e-02  1.018e-02   1.860   0.0631 .
as.factor(year)2010                            1.864e-02  9.522e-03   1.957   0.0505 .
as.factor(year)2011                            2.236e-02  9.502e-03   2.354   0.0187 *
as.factor(year)2012                            1.608e-02  9.132e-03   1.761   0.0784 .
as.factor(year)2013                            2.080e-02  9.021e-03   2.306   0.0212 *
as.factor(year)2014                            2.045e-02  8.942e-03   2.287   0.0223 *
as.factor(year)2015                            1.767e-02  8.898e-03   1.986   0.0472 *
as.factor(year)2016                            1.934e-02  8.956e-03   2.159   0.0310 *
as.factor(year)2017                            1.366e-02  8.839e-03   1.546   0.1223
as.factor(year)2018                            1.609e-02  8.740e-03   1.841   0.0659 .
as.factor(year)2019                            1.428e-02  8.662e-03   1.649   0.0994 .
as.factor(year)2020                            1.308e-02  9.815e-03   1.333   0.1829
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.04113 on 1634 degrees of freedom
Multiple R-squared:  0.02415,   Adjusted R-squared:  0.005041
F-statistic: 1.264 on 32 and 1634 DF,  p-value: 0.1485
```

```
Call:
lm(formula = CAR5 ~ greenwashing_score + log(1 + SIZE) + ROA +
    MOMENTUM + ESGscore + as.factor(Industry) + as.factor(year),
    data = merged_df)

Residuals:
     Min       1Q   Median       3Q      Max
-0.22855 -0.02036  0.00081  0.02028  0.22575

Coefficients:
                                                Estimate Std. Error t value Pr(>|t|)
(Intercept)                                   -0.0210428  0.0142080  -1.481   0.1388
greenwashing_score                             0.0005079  0.0186926   0.027   0.9783
log(1 + SIZE)                                  0.0007061  0.0008347   0.846   0.3977
ROA                                           -0.0002114  0.0001752  -1.207   0.2277
MOMENTUM                                      -0.0032468  0.0020997  -1.546   0.1222
ESGscore                                       0.0001355  0.0004745   0.286   0.7753
as.factor(Industry)Computers                  -0.0048055  0.0057592  -0.834   0.4042
as.factor(Industry)Durable manufacturers       0.0027185  0.0054005   0.503   0.6148
as.factor(Industry)Extractive industries       0.0041785  0.0073140   0.571   0.5679
as.factor(Industry)Financial institutions     -0.0132571  0.0062091  -2.135   0.0329 *
as.factor(Industry)Food                       -0.0075365  0.0065249  -1.155   0.2482
as.factor(Industry)Insurance and real estate  -0.0001463  0.0077731  -0.019   0.9850
as.factor(Industry)Mining and Construction    -0.0108419  0.0095819  -1.131   0.2580
as.factor(Industry)Pharmaceuticals             0.0007784  0.0066143   0.118   0.9063
as.factor(Industry)Retail                     -0.0004391  0.0058418  -0.075   0.9401
as.factor(Industry)Services                    0.0012757  0.0068658   0.186   0.8526
as.factor(Industry)Textiles, printing and publishing 0.0001033 0.0070970 0.015 0.9884
as.factor(Industry)Transportation             -0.0010237  0.0065234  -0.157   0.8753
as.factor(Industry)Utilities                  -0.0084765  0.0070926  -1.195   0.2322
as.factor(year)2008                            0.0259232  0.0103188   2.512   0.0121 *
as.factor(year)2009                            0.0188001  0.0101293   1.856   0.0636 .
as.factor(year)2010                            0.0186224  0.0094726   1.966   0.0495 *
as.factor(year)2011                            0.0223355  0.0094534   2.363   0.0183 *
as.factor(year)2012                            0.0160527  0.0090847   1.767   0.0774 .
as.factor(year)2013                            0.0206530  0.0089768   2.301   0.0215 *
as.factor(year)2014                            0.0189038  0.0089084   2.122   0.0340 *
as.factor(year)2015                            0.0175213  0.0088673   1.976   0.0483 *
as.factor(year)2016                            0.0193065  0.0089123   2.166   0.0304 *
as.factor(year)2017                            0.0135383  0.0087999   1.538   0.1241
as.factor(year)2018                            0.0157466  0.0087068   1.809   0.0707 .
as.factor(year)2019                            0.0140882  0.0086344   1.632   0.1030
as.factor(year)2020                            0.0127295  0.0097645   1.304   0.1925
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.04091 on 1626 degrees of freedom
  (9 observations deleted due to missingness)
Multiple R-squared:  0.02434,   Adjusted R-squared:  0.005737
F-statistic: 1.308 on 31 and 1626 DF,  p-value: 0.1199
```

○ CAR60 (GPT vs. NLP)

```
Call:
lm(formula = CAR60 ~ GWS_gpt + log(1 + SIZE) + ROA + MOMENTUM +
    ESGscore + as.factor(Industry) + as.factor(year), data = merged_df)

Residuals:
     Min       1Q   Median       3Q      Max
-0.55614 -0.04936  0.00052  0.04758  0.61531

Coefficients:
                                                 Estimate Std. Error t value Pr(>|t|)
(Intercept)                                    -0.0294165  0.0376568  -0.781    0.435
GWS_gpt                                        -0.0088963  0.0348223  -0.255    0.798
log(1 + SIZE)                                   0.0014514  0.0023461   0.619    0.536
ROA                                            -0.0001436  0.0004954  -0.290    0.772
MOMENTUM                                       -0.0356367  0.0059327  -6.007 2.33e-09 ***
ESGscore                                       -0.0005199  0.0013461  -0.386    0.699
as.factor(Industry)Computers                    0.0036924  0.0159696   0.231    0.817
as.factor(Industry)Durable manufacturers        0.0117449  0.0152171   0.772    0.440
as.factor(Industry)Extractive industries        0.0081493  0.0207201   0.393    0.694
as.factor(Industry)Financial institutions       0.0126997  0.0175840   0.722    0.470
as.factor(Industry)Food                         0.0102975  0.0178815   0.576    0.565
as.factor(Industry)Insurance and real estate    0.0152373  0.0214489   0.710    0.478
as.factor(Industry)Mining and Construction      0.0225508  0.0270605   0.833    0.405
as.factor(Industry)Others                      -0.0924198  0.0836637  -1.105    0.269
as.factor(Industry)Pharmaceuticals              0.0242791  0.0187308   1.296    0.195
as.factor(Industry)Retail                       0.0234599  0.0163634   1.434    0.152
as.factor(Industry)Services                     0.0088173  0.0192608   0.458    0.647
as.factor(Industry)Textiles, printing and publishing  0.0193015  0.0200116   0.965    0.335
as.factor(Industry)Transportation              -0.0085239  0.0184801  -0.461    0.645
as.factor(Industry)Utilities                    0.0013181  0.0194063   0.068    0.946
as.factor(year)2008                             0.0254834  0.0292211   0.872    0.383
as.factor(year)2009                             0.0090922  0.0287126   0.317    0.752
as.factor(year)2010                            -0.0029384  0.0268460  -0.109    0.913
as.factor(year)2011                             0.0017891  0.0267885   0.067    0.947
as.factor(year)2012                             0.0272947  0.0257459   1.060    0.289
as.factor(year)2013                            -0.0037012  0.0254319  -0.146    0.884
as.factor(year)2014                             0.0279332  0.0252104   1.108    0.268
as.factor(year)2015                             0.0220172  0.0250854   0.878    0.380
as.factor(year)2016                             0.0072015  0.0252503   0.285    0.776
as.factor(year)2017                             0.0016574  0.0249186   0.067    0.947
as.factor(year)2018                             0.0041968  0.0246395   0.170    0.865
as.factor(year)2019                             0.0087699  0.0244212   0.359    0.720
as.factor(year)2020                             0.0101871  0.0276722   0.368    0.713
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.116 on 1634 degrees of freedom
Multiple R-squared:  0.03652,	Adjusted R-squared:  0.01765
F-statistic: 1.935 on 32 and 1634 DF,  p-value: 0.001363
```

```
Call:
lm(formula = CAR60 ~ greenwashing_score + log(1 + SIZE) + ROA +
    MOMENTUM + ESGscore + as.factor(Industry) + as.factor(year),
    data = merged_df)

Residuals:
     Min       1Q   Median       3Q      Max
-0.55415 -0.04968  0.00025  0.04832  0.61443

Coefficients:
                                                 Estimate Std. Error t value Pr(>|t|)
(Intercept)                                    -0.0349489  0.0402515  -0.868    0.385
greenwashing_score                             -0.0043488  0.0529563  -0.082    0.935
log(1 + SIZE)                                   0.0018920  0.0023646   0.800    0.424
ROA                                            -0.0001348  0.0004964  -0.271    0.786
MOMENTUM                                       -0.0365955  0.0059486  -6.152 9.62e-10 ***
ESGscore                                       -0.0004544  0.0013441  -0.338    0.735
as.factor(Industry)Computers                    0.0030420  0.0163159   0.186    0.852
as.factor(Industry)Durable manufacturers        0.0115130  0.0152996   0.753    0.452
as.factor(Industry)Extractive industries        0.0074474  0.0207207   0.359    0.719
as.factor(Industry)Financial institutions       0.0112529  0.0175904   0.640    0.522
as.factor(Industry)Food                         0.0101270  0.0184851   0.548    0.584
as.factor(Industry)Insurance and real estate    0.0155475  0.0220212   0.706    0.480
as.factor(Industry)Mining and Construction      0.0227385  0.0271458   0.838    0.402
as.factor(Industry)Pharmaceuticals              0.0236892  0.0187384   1.264    0.206
as.factor(Industry)Retail                       0.0220110  0.0165499   1.330    0.184
as.factor(Industry)Services                     0.0038632  0.0194509   0.199    0.843
as.factor(Industry)Textiles, printing and publishing  0.0192326  0.0201059   0.957    0.339
as.factor(Industry)Transportation              -0.0093701  0.0184809  -0.507    0.612
as.factor(Industry)Utilities                    0.0001740  0.0200933   0.009    0.993
as.factor(year)2008                             0.0255452  0.0292334   0.874    0.382
as.factor(year)2009                             0.0087934  0.0286963   0.306    0.759
as.factor(year)2010                            -0.0030573  0.0268360  -0.114    0.909
as.factor(year)2011                             0.0016897  0.0267815   0.063    0.950
as.factor(year)2012                             0.0275312  0.0257370   1.070    0.285
as.factor(year)2013                            -0.0035219  0.0254315  -0.138    0.890
as.factor(year)2014                             0.0273736  0.0252375   1.085    0.278
as.factor(year)2015                             0.0195272  0.0251211   0.777    0.437
as.factor(year)2016                             0.0075566  0.0252487   0.299    0.765
as.factor(year)2017                             0.0017679  0.0249302   0.071    0.943
as.factor(year)2018                             0.0034856  0.0246666   0.141    0.888
as.factor(year)2019                             0.0090248  0.0244615   0.369    0.712
as.factor(year)2020                             0.0102335  0.0276630   0.370    0.711
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1159 on 1626 degrees of freedom
  (9 observations deleted due to missingness)
Multiple R-squared:  0.0368,	Adjusted R-squared:  0.01843
F-statistic: 2.004 on 31 and 1626 DF,  p-value: 0.0008952
```

● Appendix 10 - Model 2 Regressions (Section 4.3)

○ CAR2 (GPT vs. NLP)

```
Call:
lm(formula = CAR2 ~ GWS_gpt * ESGscore + log(1 + SIZE) + ROA +
    MOMENTUM + as.factor(Industry) + as.factor(year), data = merged_df)

Residuals:
      Min        1Q    Median        3Q       Max
-0.184312 -0.012900 -0.000193  0.012319  0.213204

Coefficients:
                                                Estimate Std. Error t value Pr(>|t|)
(Intercept)                                   -5.121e-03  1.326e-02  -0.386  0.69940
GWS_gpt                                       -1.147e-02  2.426e-02  -0.473  0.63637
ESGscore                                      -8.042e-04  1.727e-03  -0.466  0.64155
log(1 + SIZE)                                 -4.006e-04  5.937e-04  -0.675  0.49989
ROA                                            5.293e-05  1.254e-04   0.422  0.67291
MOMENTUM                                      -3.314e-03  1.501e-03  -2.208  0.02735 *
as.factor(Industry)Computers                   4.544e-03  4.042e-03   1.124  0.26117
as.factor(Industry)Durable manufacturers       8.216e-03  3.849e-03   2.135  0.03295 *
as.factor(Industry)Extractive industries       6.213e-03  5.241e-03   1.185  0.23605
as.factor(Industry)Financial institutions     -1.074e-03  4.449e-03  -0.241  0.80928
as.factor(Industry)Food                       -1.959e-03  4.524e-03  -0.433  0.66513
as.factor(Industry)Insurance and real estate  -1.556e-06  5.425e-03   0.000  0.99977
as.factor(Industry)Mining and Construction    -7.082e-03  6.845e-03  -1.035  0.30102
as.factor(Industry)Others                      2.654e-03  2.117e-02   0.125  0.90028
as.factor(Industry)Pharmaceuticals             5.312e-03  4.743e-03   1.120  0.26290
as.factor(Industry)Retail                      7.224e-03  4.139e-03   1.745  0.08109 .
as.factor(Industry)Services                    6.946e-04  4.872e-03   0.143  0.88664
as.factor(Industry)Textiles, printing and publishing  3.120e-03  5.062e-03   0.616  0.53767
as.factor(Industry)Transportation              2.547e-03  4.675e-03   0.545  0.58606
as.factor(Industry)Utilities                   2.616e-03  4.909e-03   0.533  0.59415
as.factor(year)2008                            2.094e-02  7.393e-03   2.832  0.00468 **
as.factor(year)2009                            1.641e-02  7.264e-03   2.259  0.02404 *
as.factor(year)2010                            1.266e-02  6.790e-03   1.865  0.06241 .
as.factor(year)2011                            1.750e-02  6.777e-03   2.583  0.00989 **
as.factor(year)2012                            1.002e-02  6.515e-03   1.539  0.12407
as.factor(year)2013                            1.412e-02  6.436e-03   2.193  0.02843 *
as.factor(year)2014                            1.253e-02  6.379e-03   1.965  0.04964 *
as.factor(year)2015                            1.506e-02  6.349e-03   2.372  0.01780 *
as.factor(year)2016                            1.279e-02  6.389e-03   2.002  0.04544 *
as.factor(year)2017                            1.070e-02  6.306e-03   1.697  0.08979 .
as.factor(year)2018                            1.280e-02  6.235e-03   2.054  0.04016 *
as.factor(year)2019                            1.021e-02  6.178e-03   1.652  0.09874 .
as.factor(year)2020                            1.277e-02  7.004e-03   1.823  0.06845 .
GWS_gpt:ESGscore                               1.179e-03  4.205e-03   0.280  0.77920
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02933 on 1633 degrees of freedom
Multiple R-squared:  0.0306,	Adjusted R-squared:  0.01101
F-statistic: 1.562 on 33 and 1633 DF,  p-value: 0.02237
```

```
Call:
lm(formula = CAR2 ~ greenwashing_score * ESGscore + log(1 + SIZE) +
    ROA + MOMENTUM + as.factor(Industry) + as.factor(year), data = merged_df)

Residuals:
     Min       1Q   Median       3Q      Max
-0.18437 -0.01281 -0.00003  0.01243  0.21259

Coefficients:
                                                Estimate Std. Error t value Pr(>|t|)
(Intercept)                                   -1.576e-02  1.650e-02  -0.955  0.33983
greenwashing_score                             1.118e-02  2.986e-02   0.374  0.70828
ESGscore                                       5.103e-04  2.435e-03   0.210  0.83404
log(1 + SIZE)                                 -3.089e-04  5.975e-04  -0.517  0.60520
ROA                                            4.925e-05  1.250e-04   0.394  0.69368
MOMENTUM                                      -3.293e-03  1.498e-03  -2.198  0.02810 *
as.factor(Industry)Computers                   4.023e-03  4.113e-03   0.978  0.32820
as.factor(Industry)Durable manufacturers       7.919e-03  3.863e-03   2.050  0.04053 *
as.factor(Industry)Extractive industries       5.838e-03  5.238e-03   1.115  0.26519
as.factor(Industry)Financial institutions     -1.550e-03  4.441e-03  -0.349  0.72708
as.factor(Industry)Food                       -2.328e-03  4.656e-03  -0.500  0.61707
as.factor(Industry)Insurance and real estate  -3.249e-04  5.547e-03  -0.059  0.95330
as.factor(Industry)Mining and Construction    -7.384e-03  6.838e-03  -1.080  0.28034
as.factor(Industry)Pharmaceuticals             5.185e-03  4.726e-03   1.097  0.27278
as.factor(Industry)Retail                      7.107e-03  4.180e-03   1.700  0.08926 .
as.factor(Industry)Services                    7.856e-04  4.902e-03   0.160  0.87270
as.factor(Industry)Textiles, printing and publishing  2.776e-03  5.069e-03   0.548  0.58408
as.factor(Industry)Transportation              2.320e-03  4.662e-03   0.498  0.61880
as.factor(Industry)Utilities                   3.770e-04  5.067e-03   0.074  0.94070
as.factor(year)2008                            2.084e-02  7.363e-03   2.831  0.00470 **
as.factor(year)2009                            1.637e-02  7.228e-03   2.264  0.02368 *
as.factor(year)2010                            1.265e-02  6.760e-03   1.871  0.06146 .
as.factor(year)2011                            1.757e-02  6.748e-03   2.603  0.00932 **
as.factor(year)2012                            1.012e-02  6.487e-03   1.560  0.11896
as.factor(year)2013                            1.407e-02  6.407e-03   2.196  0.02820 *
as.factor(year)2014                            1.150e-02  6.357e-03   1.810  0.07054 .
as.factor(year)2015                            1.504e-02  6.329e-03   2.377  0.01759 *
as.factor(year)2016                            1.276e-02  6.361e-03   2.006  0.04504 *
as.factor(year)2017                            1.067e-02  6.281e-03   1.700  0.08942 .
as.factor(year)2018                            1.270e-02  6.213e-03   2.045  0.04105 *
as.factor(year)2019                            1.010e-02  6.163e-03   1.640  0.10127
as.factor(year)2020                            1.249e-02  6.968e-03   1.793  0.07316 .
greenwashing_score:ESGscore                   -1.649e-03  5.050e-03  -0.327  0.74401
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02919 on 1625 degrees of freedom
  (9 observations deleted due to missingness)
Multiple R-squared:  0.03109,	Adjusted R-squared:  0.01201
F-statistic:  1.63 on 32 and 1625 DF,  p-value: 0.01483
```

○ CAR5 (GPT vs. NLP)

```
Call:
lm(formula = CAR5 ~ GWS_gpt * ESGscore + log(1 + SIZE) + ROA +
    MOMENTUM + as.factor(Industry) + as.factor(year), data = merged_df)

Residuals:
      Min        1Q    Median        3Q       Max
-0.228554 -0.020418  0.000742  0.019911  0.225906

Coefficients:
                                                  Estimate Std. Error t value Pr(>|t|)
(Intercept)                                     -1.079e-02  1.860e-02  -0.580   0.5618
GWS_gpt                                         -2.186e-02  3.403e-02  -0.643   0.5206
ESGscore                                        -1.083e-03  2.423e-03  -0.447   0.6550
log(1 + SIZE)                                    5.878e-04  8.327e-04   0.706   0.4804
ROA                                             -2.050e-04  1.758e-04  -1.166   0.2438
MOMENTUM                                        -3.140e-03  2.105e-03  -1.492   0.1359
as.factor(Industry)Computers                    -4.277e-03  5.670e-03  -0.754   0.4508
as.factor(Industry)Durable manufacturers         2.938e-03  5.399e-03   0.544   0.5864
as.factor(Industry)Extractive industries         4.528e-03  7.351e-03   0.616   0.5380
as.factor(Industry)Financial institutions       -1.276e-02  6.241e-03  -2.045   0.0410 *
as.factor(Industry)Food                         -7.301e-03  6.346e-03  -1.150   0.2501
as.factor(Industry)Insurance and real estate    -5.677e-05  7.610e-03  -0.007   0.9940
as.factor(Industry)Mining and Construction      -1.061e-02  9.601e-03  -1.105   0.2693
as.factor(Industry)Others                       -3.485e-02  2.970e-02  -1.173   0.2408
as.factor(Industry)Pharmaceuticals               7.677e-04  6.652e-03   0.115   0.9081
as.factor(Industry)Retail                       -6.884e-05  5.805e-03  -0.012   0.9905
as.factor(Industry)Services                      8.872e-04  6.834e-03   0.130   0.8967
as.factor(Industry)Textiles, printing and publishing  3.816e-04  7.100e-03   0.054   0.9571
as.factor(Industry)Transportation               -8.078e-04  6.558e-03  -0.123   0.9020
as.factor(Industry)Utilities                    -5.752e-03  6.885e-03  -0.835   0.4036
as.factor(year)2008                              2.603e-02  1.037e-02   2.510   0.0122 *
as.factor(year)2009                              1.883e-02  1.019e-02   1.848   0.0648 .
as.factor(year)2010                              1.865e-02  9.524e-03   1.958   0.0504 .
as.factor(year)2011                              2.227e-02  9.506e-03   2.343   0.0193 *
as.factor(year)2012                              1.595e-02  9.138e-03   1.745   0.0811 .
as.factor(year)2013                              2.067e-02  9.027e-03   2.290   0.0221 *
as.factor(year)2014                              2.035e-02  8.947e-03   2.274   0.0231 *
as.factor(year)2015                              1.752e-02  8.906e-03   1.967   0.0494 *
as.factor(year)2016                              1.922e-02  8.962e-03   2.145   0.0321 *
as.factor(year)2017                              1.354e-02  8.845e-03   1.531   0.1260
as.factor(year)2018                              1.597e-02  8.745e-03   1.826   0.0680 .
as.factor(year)2019                              1.422e-02  8.665e-03   1.641   0.1011
as.factor(year)2020                              1.290e-02  9.825e-03   1.313   0.1892
GWS_gpt:ESGscore                                 2.805e-03  5.899e-03   0.476   0.6345
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.04114 on 1633 degrees of freedom
Multiple R-squared:  0.02429,	Adjusted R-squared:  0.00457
F-statistic: 1.232 on 33 and 1633 DF,  p-value: 0.1723
```

```
Call:
lm(formula = CAR5 ~ greenwashing_score * ESGscore + log(1 + SIZE) +
    ROA + MOMENTUM + as.factor(Industry) + as.factor(year), data = merged_df)

Residuals:
      Min        1Q    Median        3Q       Max
-0.228699 -0.020742  0.000862  0.020181  0.224162

Coefficients:
                                                  Estimate Std. Error t value Pr(>|t|)
(Intercept)                                      8.732e-03  2.312e-02   0.378   0.7057
greenwashing_score                              -6.058e-02  4.183e-02  -1.448   0.1477
ESGscore                                        -5.378e-03  3.411e-03  -1.577   0.1150
log(1 + SIZE)                                    5.955e-04  8.370e-04   0.711   0.4769
ROA                                             -2.084e-04  1.751e-04  -1.190   0.2342
MOMENTUM                                        -3.224e-03  2.099e-03  -1.536   0.1247
as.factor(Industry)Computers                    -4.420e-03  5.761e-03  -0.767   0.4430
as.factor(Industry)Durable manufacturers         3.344e-03  5.411e-03   0.618   0.5366
as.factor(Industry)Extractive industries         5.208e-03  7.337e-03   0.710   0.4780
as.factor(Industry)Financial institutions       -1.255e-02  6.221e-03  -2.017   0.0439 *
as.factor(Industry)Food                         -7.429e-03  6.522e-03  -1.139   0.2548
as.factor(Industry)Insurance and real estate     7.396e-05  7.770e-03   0.010   0.9924
as.factor(Industry)Mining and Construction      -1.061e-02  9.578e-03  -1.108   0.2681
as.factor(Industry)Pharmaceuticals               1.364e-03  6.621e-03   0.206   0.8368
as.factor(Industry)Retail                        2.765e-04  5.855e-03   0.047   0.9623
as.factor(Industry)Services                      1.669e-03  6.867e-03   0.243   0.8080
as.factor(Industry)Textiles, printing and publishing  6.314e-04  7.101e-03   0.089   0.9292
as.factor(Industry)Transportation               -4.104e-04  6.531e-03  -0.063   0.9499
as.factor(Industry)Utilities                    -7.908e-03  7.097e-03  -1.114   0.2654
as.factor(year)2008                              2.593e-02  1.031e-02   2.514   0.0120 *
as.factor(year)2009                              1.869e-02  1.012e-02   1.846   0.0651 .
as.factor(year)2010                              1.833e-02  9.469e-03   1.935   0.0531 .
as.factor(year)2011                              2.190e-02  9.452e-03   2.317   0.0206 *
as.factor(year)2012                              1.550e-02  9.086e-03   1.706   0.0883 .
as.factor(year)2013                              2.034e-02  8.974e-03   2.266   0.0236 *
as.factor(year)2014                              1.877e-02  8.904e-03   2.108   0.0351 *
as.factor(year)2015                              1.717e-02  8.865e-03   1.937   0.0529 .
as.factor(year)2016                              1.893e-02  8.911e-03   2.124   0.0338 *
as.factor(year)2017                              1.322e-02  8.798e-03   1.503   0.1330
as.factor(year)2018                              1.562e-02  8.703e-03   1.795   0.0729 .
as.factor(year)2019                              1.376e-02  8.632e-03   1.594   0.1112
as.factor(year)2020                              1.259e-02  9.760e-03   1.290   0.1974
greenwashing_score:ESGscore                      1.155e-02  7.074e-03   1.632   0.1028
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.04088 on 1625 degrees of freedom
  (9 observations deleted due to missingness)
Multiple R-squared:  0.02594,	Adjusted R-squared:  0.006754
F-statistic: 1.352 on 32 and 1625 DF,  p-value: 0.09111
```

○ CAR60 (GPT vs. NLP)

```
Call:
lm(formula = CAR60 ~ GWS_gpt * ESGscore + log(1 + SIZE) + ROA +
    MOMENTUM + as.factor(Industry) + as.factor(year), data = merged_df)

Residuals:
     Min       1Q   Median       3Q      Max
-0.55626 -0.04924  0.00055  0.04743  0.61529

Coefficients:
                                                  Estimate Std. Error t value Pr(>|t|)
(Intercept)                                     -0.0324833  0.0524388  -0.619    0.536
GWS_gpt                                         -0.0013820  0.0959342  -0.014    0.989
ESGscore                                         0.0000431  0.0068308   0.006    0.995
log(1 + SIZE)                                    0.0014457  0.0023477   0.616    0.538
ROA                                             -0.0001425  0.0004957  -0.288    0.774
MOMENTUM                                        -0.0356368  0.0059345  -6.005 2.35e-09 ***
as.factor(Industry)Computers                     0.0037454  0.0159869   0.234    0.815
as.factor(Industry)Durable manufacturers         0.0117320  0.0152225   0.771    0.441
as.factor(Industry)Extractive industries         0.0081312  0.0207275   0.392    0.695
as.factor(Industry)Financial institutions        0.0127402  0.0175959   0.724    0.469
as.factor(Industry)Food                          0.0103348  0.0178924   0.578    0.564
as.factor(Industry)Insurance and real estate     0.0152508  0.0214561   0.711    0.477
as.factor(Industry)Mining and Construction       0.0225247  0.0270704   0.832    0.405
as.factor(Industry)Others                       -0.0921748  0.0837399  -1.101    0.271
as.factor(Industry)Pharmaceuticals               0.0243511  0.0187561   1.298    0.194
as.factor(Industry)Retail                        0.0234565  0.0163684   1.433    0.152
as.factor(Industry)Services                      0.0088011  0.0192676   0.457    0.648
as.factor(Industry)Textiles, printing and publishing  0.0192914  0.0200180   0.964    0.335
as.factor(Industry)Transportation               -0.0084901  0.0184900  -0.459    0.646
as.factor(Industry)Utilities                     0.0013154  0.0194122   0.068    0.946
as.factor(year)2008                              0.0254321  0.0292363   0.870    0.384
as.factor(year)2009                              0.0091478  0.0287289   0.318    0.750
as.factor(year)2010                             -0.0029423  0.0268542  -0.110    0.913
as.factor(year)2011                              0.0018352  0.0268022   0.068    0.945
as.factor(year)2012                              0.0273605  0.0257656   1.062    0.288
as.factor(year)2013                             -0.0036367  0.0254512  -0.143    0.886
as.factor(year)2014                              0.0279859  0.0252258   1.109    0.267
as.factor(year)2015                              0.0220939  0.0251096   0.880    0.379
as.factor(year)2016                              0.0072603  0.0252676   0.287    0.774
as.factor(year)2017                              0.0017203  0.0249374   0.069    0.945
as.factor(year)2018                              0.0042550  0.0246568   0.173    0.863
as.factor(year)2019                              0.0088033  0.0244318   0.360    0.719
as.factor(year)2020                              0.0102748  0.0277003   0.371    0.711
GWS_gpt:ESGscore                                -0.0013981  0.0166308  -0.084    0.933
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.116 on 1633 degrees of freedom
Multiple R-squared:  0.03652,	Adjusted R-squared:  0.01705
F-statistic: 1.876 on 33 and 1633 DF,  p-value: 0.001961
```

```
Call:
lm(formula = CAR60 ~ greenwashing_score * ESGscore + log(1 +
    SIZE) + ROA + MOMENTUM + as.factor(Industry) + as.factor(year),
    data = merged_df)

Residuals:
     Min       1Q   Median       3Q      Max
-0.55400 -0.04932  0.00030  0.04824  0.61509

Coefficients:
                                                  Estimate Std. Error t value Pr(>|t|)
(Intercept)                                     -0.0529871  0.0655396  -0.808    0.419
greenwashing_score                               0.0326619  0.1185941   0.275    0.783
ESGscore                                         0.0028858  0.0096702   0.298    0.765
log(1 + SIZE)                                    0.0019589  0.0023730   0.826    0.409
ROA                                             -0.0001366  0.0004965  -0.275    0.783
MOMENTUM                                        -0.0366094  0.0059503  -6.152 9.59e-10 ***
as.factor(Industry)Computers                     0.0028086  0.0163341   0.172    0.864
as.factor(Industry)Durable manufacturers         0.0111339  0.0153423   0.726    0.468
as.factor(Industry)Extractive industries         0.0068240  0.0208032   0.328    0.743
as.factor(Industry)Financial institutions        0.0108230  0.0176382   0.614    0.540
as.factor(Industry)Food                          0.0100619  0.0184911   0.544    0.586
as.factor(Industry)Insurance and real estate     0.0154141  0.0220305   0.700    0.484
as.factor(Industry)Mining and Construction       0.0225992  0.0271561   0.832    0.405
as.factor(Industry)Pharmaceuticals               0.0233345  0.0187710   1.243    0.214
as.factor(Industry)Retail                        0.0215775  0.0166009   1.300    0.194
as.factor(Industry)Services                      0.0036252  0.0194681   0.186    0.852
as.factor(Industry)Textiles, printing and publishing  0.0189127  0.0201322   0.939    0.348
as.factor(Industry)Transportation               -0.0097416  0.0185166  -0.526    0.599
as.factor(Industry)Utilities                    -0.0001707  0.0201230  -0.008    0.993
as.factor(year)2008                              0.0255433  0.0292413   0.874    0.383
as.factor(year)2009                              0.0088617  0.0287048   0.309    0.758
as.factor(year)2010                             -0.0028786  0.0268482  -0.107    0.915
as.factor(year)2011                              0.0019511  0.0267993   0.073    0.942
as.factor(year)2012                              0.0278669  0.0257619   1.082    0.280
as.factor(year)2013                             -0.0033311  0.0254442  -0.131    0.896
as.factor(year)2014                              0.0274522  0.0252454   1.087    0.277
as.factor(year)2015                              0.0197399  0.0251353   0.785    0.432
as.factor(year)2016                              0.0077850  0.0252640   0.308    0.758
as.factor(year)2017                              0.0019592  0.0249430   0.079    0.937
as.factor(year)2018                              0.0035636  0.0246743   0.144    0.885
as.factor(year)2019                              0.0092258  0.0244749   0.377    0.706
as.factor(year)2020                              0.0103198  0.0276716   0.373    0.709
greenwashing_score:ESGscore                     -0.0069953  0.0200550  -0.349    0.727
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1159 on 1625 degrees of freedom
  (9 observations deleted due to missingness)
Multiple R-squared:  0.03687,	Adjusted R-squared:  0.0179
F-statistic: 1.944 on 32 and 1625 DF,  p-value: 0.00127
```

- **Appendix 11 - Model 3 Regressions (Section 4.4)**
    - CAR2 (GPT vs. NLP)

```
lm(formula = CAR2 ~ sentiment_score_gpt + specificity_score_gpt +
    buzzwords_score_gpt + hedging_score_gpt + candor_score_gpt +
    log(1 + SIZE) + ROA + MOMENTUM + ESGscore + as.factor(Industry) +
    as.factor(year), data = merged_df)

Residuals:
      Min        1Q    Median        3Q       Max
-0.184030 -0.012844 -0.000173  0.012470  0.211558

Coefficients:
                                                 Estimate Std. Error t value Pr(>|t|)
(Intercept)                                    -1.578e-02  1.007e-02  -1.566  0.11744
sentiment_score_gpt                            -5.842e-03  1.770e-02  -0.330  0.74140
specificity_score_gpt                           3.474e-03  3.463e-03   1.003  0.31588
buzzwords_score_gpt                             8.846e-03  9.802e-03   0.902  0.36696
hedging_score_gpt                               3.911e-03  1.602e-02   0.244  0.80721
candor_score_gpt                                7.348e-04  6.343e-03   0.116  0.90779
log(1 + SIZE)                                  -5.017e-04  6.020e-04  -0.833  0.40469
ROA                                             5.585e-05  1.254e-04   0.445  0.65617
MOMENTUM                                       -3.238e-03  1.502e-03  -2.155  0.03128 *
ESGscore                                       -3.605e-04  3.440e-04  -1.048  0.29488
as.factor(Industry)Computers                    4.970e-03  4.052e-03   1.227  0.22016
as.factor(Industry)Durable manufacturers        8.507e-03  3.855e-03   2.207  0.02746 *
as.factor(Industry)Extractive industries        6.115e-03  5.284e-03   1.157  0.24740
as.factor(Industry)Financial institutions      -3.403e-04  4.509e-03  -0.075  0.93985
as.factor(Industry)Food                        -1.509e-03  4.539e-03  -0.332  0.73965
as.factor(Industry)Insurance and real estate    2.946e-04  5.437e-03   0.054  0.95680
as.factor(Industry)Mining and Construction     -6.780e-03  6.851e-03  -0.990  0.32248
as.factor(Industry)Others                       3.980e-03  2.120e-02   0.188  0.85115
as.factor(Industry)Pharmaceuticals              5.816e-03  4.778e-03   1.217  0.22374
as.factor(Industry)Retail                       7.890e-03  4.179e-03   1.888  0.05920 .
as.factor(Industry)Services                     1.468e-03  4.899e-03   0.300  0.76452
as.factor(Industry)Textiles, printing and publishing 3.470e-03  5.070e-03   0.684  0.49382
as.factor(Industry)Transportation               3.107e-03  4.691e-03   0.662  0.50784
as.factor(Industry)Utilities                    2.813e-03  4.916e-03   0.572  0.56731
as.factor(year)2008                             2.103e-02  7.398e-03   2.843  0.00453 **
as.factor(year)2009                             1.598e-02  7.270e-03   2.198  0.02809 *
as.factor(year)2010                             1.240e-02  6.792e-03   1.826  0.06799 .
as.factor(year)2011                             1.726e-02  6.781e-03   2.545  0.01103 *
as.factor(year)2012                             1.005e-02  6.517e-03   1.542  0.12320
as.factor(year)2013                             1.381e-02  6.442e-03   2.144  0.03215 *
as.factor(year)2014                             1.226e-02  6.391e-03   1.918  0.05529 .
as.factor(year)2015                             1.502e-02  6.352e-03   2.365  0.01815 *
as.factor(year)2016                             1.248e-02  6.399e-03   1.951  0.05121 .
as.factor(year)2017                             1.043e-02  6.314e-03   1.652  0.09882 .
as.factor(year)2018                             1.248e-02  6.244e-03   1.998  0.04589 *
as.factor(year)2019                             9.776e-03  6.210e-03   1.574  0.11565
as.factor(year)2020                             1.257e-02  7.027e-03   1.789  0.07383 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.02933 on 1630 degrees of freedom
Multiple R-squared:  0.03233,	Adjusted R-squared:  0.01095
F-statistic: 1.513 on 36 and 1630 DF,  p-value: 0.02673
```

```
Call:
lm(formula = CAR2 ~ sentiment_risk_score + specificity_score +
    buzzwords_score_industry_norm + hedging_density_norma + candor_risk_component +
    log(1 + SIZE) + ROA + MOMENTUM + ESGscore + as.factor(Industry) +
    as.factor(year), data = merged_df)

Residuals:
      Min        1Q    Median        3Q       Max
-0.184275 -0.013192  0.000019  0.012919  0.208912

Coefficients:
                                                 Estimate Std. Error t value Pr(>|t|)
(Intercept)                                    -4.054e-02  1.665e-02  -2.435  0.01502 *
sentiment_risk_score                           -1.386e-02  1.063e-02  -1.304  0.19253
specificity_score                               1.969e-02  1.117e-02   1.764  0.07797 .
buzzwords_score_industry_norm                   6.606e-03  5.857e-03   1.128  0.25953
hedging_density_norma                           1.645e-02  7.973e-03   2.063  0.03923 *
candor_risk_component                           1.201e-02  1.040e-02   1.155  0.24834
log(1 + SIZE)                                  -3.558e-04  5.948e-04  -0.598  0.54980
ROA                                             5.796e-05  1.251e-04   0.463  0.64317
MOMENTUM                                       -3.317e-03  1.495e-03  -2.218  0.02667 *
ESGscore                                       -3.811e-04  3.413e-04  -1.117  0.26435
as.factor(Industry)Computers                    3.149e-03  4.171e-03   0.755  0.45047
as.factor(Industry)Durable manufacturers        7.588e-03  3.867e-03   1.962  0.04990 *
as.factor(Industry)Extractive industries        5.746e-03  5.225e-03   1.100  0.27159
as.factor(Industry)Financial institutions      -1.124e-03  4.432e-03  -0.253  0.79992
as.factor(Industry)Food                        -3.741e-03  4.825e-03  -0.775  0.43820
as.factor(Industry)Insurance and real estate   -2.206e-03  5.713e-03  -0.386  0.69941
as.factor(Industry)Mining and Construction     -8.148e-03  6.848e-03  -1.190  0.23427
as.factor(Industry)Pharmaceuticals              4.978e-03  4.716e-03   1.056  0.29130
as.factor(Industry)Retail                       8.914e-03  4.255e-03   2.095  0.03635 *
as.factor(Industry)Services                     7.160e-04  4.890e-03   0.146  0.88361
as.factor(Industry)Textiles, printing and publishing  2.188e-03  5.087e-03   0.430  0.66716
as.factor(Industry)Transportation               1.966e-03  4.651e-03   0.423  0.67259
as.factor(Industry)Utilities                   -1.675e-03  5.218e-03  -0.321  0.74827
as.factor(year)2008                             2.079e-02  7.354e-03   2.827  0.00475 **
as.factor(year)2009                             1.630e-02  7.215e-03   2.259  0.02400 *
as.factor(year)2010                             1.292e-02  6.747e-03   1.915  0.05569 .
as.factor(year)2011                             1.776e-02  6.733e-03   2.638  0.00842 **
as.factor(year)2012                             1.044e-02  6.472e-03   1.613  0.10693
as.factor(year)2013                             1.402e-02  6.394e-03   2.192  0.02851 *
as.factor(year)2014                             1.126e-02  6.344e-03   1.775  0.07606 .
as.factor(year)2015                             1.517e-02  6.319e-03   2.401  0.01648 *
as.factor(year)2016                             1.244e-02  6.349e-03   1.959  0.05026 .
as.factor(year)2017                             1.065e-02  6.274e-03   1.697  0.08985 .
as.factor(year)2018                             1.250e-02  6.205e-03   2.014  0.04420 *
as.factor(year)2019                             9.567e-03  6.154e-03   1.555  0.12025
as.factor(year)2020                             1.212e-02  6.957e-03   1.742  0.08172 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.02913 on 1622 degrees of freedom
  (9 observations deleted due to missingness)
Multiple R-squared:  0.03696,	Adjusted R-squared:  0.01618
F-statistic: 1.779 on 35 and 1622 DF,  p-value: 0.003522
```

- CAR5 (GPT vs. NLP)

```
lm(formula = CAR5 ~ sentiment_score_gpt + specificity_score_gpt +
    buzzwords_score_gpt + hedging_score_gpt + candor_score_gpt +
    log(1 + SIZE) + ROA + MOMENTUM + ESGscore + as.factor(Industry) +
    as.factor(year), data = merged_df)

Residuals:
     Min       1Q   Median       3Q      Max
-0.22903 -0.02026  0.00046  0.02004  0.22567

Coefficients:
                                                 Estimate Std. Error t value Pr(>|t|)
(Intercept)                                    -2.538e-02  1.412e-02  -1.797   0.0726 .
sentiment_score_gpt                            -2.034e-02  2.482e-02  -0.820   0.4126
specificity_score_gpt                           4.043e-03  4.855e-03   0.833   0.4051
buzzwords_score_gpt                             6.360e-03  1.374e-02   0.463   0.6436
hedging_score_gpt                               3.178e-02  2.247e-02   1.414   0.1575
candor_score_gpt                               -2.076e-03  8.894e-03  -0.233   0.8155
log(1 + SIZE)                                   5.158e-04  8.441e-04   0.611   0.5412
ROA                                            -1.923e-04  1.759e-04  -1.093   0.2744
MOMENTUM                                       -3.063e-03  2.107e-03  -1.454   0.1461
ESGscore                                        2.417e-05  4.824e-04   0.050   0.9600
as.factor(Industry)Computers                   -3.817e-03  5.681e-03  -0.672   0.5017
as.factor(Industry)Durable manufacturers        3.128e-03  5.405e-03   0.579   0.5629
as.factor(Industry)Extractive industries        4.114e-03  7.410e-03   0.555   0.5788
as.factor(Industry)Financial institutions      -1.233e-02  6.322e-03  -1.950   0.0514 .
as.factor(Industry)Food                        -7.084e-03  6.364e-03  -1.113   0.2658
as.factor(Industry)Insurance and real estate   -9.980e-05  7.624e-03  -0.013   0.9896
as.factor(Industry)Mining and Construction     -1.061e-02  9.606e-03  -1.104   0.2696
as.factor(Industry)Others                      -3.183e-02  2.973e-02  -1.071   0.2845
as.factor(Industry)Pharmaceuticals              6.443e-04  6.700e-03   0.096   0.9234
as.factor(Industry)Retail                       6.867e-05  5.860e-03   0.012   0.9907
as.factor(Industry)Services                     1.539e-03  6.869e-03   0.224   0.8227
as.factor(Industry)Textiles, printing and publishing 7.196e-04  7.109e-03   0.101   0.9194
as.factor(Industry)Transportation              -4.332e-04  6.578e-03  -0.066   0.9475
as.factor(Industry)Utilities                   -5.737e-03  6.893e-03  -0.832   0.4054
as.factor(year)2008                             2.575e-02  1.037e-02   2.483   0.0131 *
as.factor(year)2009                             1.829e-02  1.019e-02   1.794   0.0730 .
as.factor(year)2010                             1.850e-02  9.524e-03   1.942   0.0523 .
as.factor(year)2011                             2.222e-02  9.509e-03   2.337   0.0196 *
as.factor(year)2012                             1.601e-02  9.138e-03   1.752   0.0800 .
as.factor(year)2013                             2.069e-02  9.032e-03   2.291   0.0221 *
as.factor(year)2014                             2.034e-02  8.962e-03   2.270   0.0234 *
as.factor(year)2015                             1.775e-02  8.907e-03   1.992   0.0465 *
as.factor(year)2016                             1.889e-02  8.972e-03   2.106   0.0354 *
as.factor(year)2017                             1.344e-02  8.854e-03   1.519   0.1291
as.factor(year)2018                             1.565e-02  8.756e-03   1.788   0.0740 .
as.factor(year)2019                             1.395e-02  8.708e-03   1.602   0.1093
as.factor(year)2020                             1.302e-02  9.853e-03   1.322   0.1864
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.04113 on 1630 degrees of freedom
Multiple R-squared:  0.02663,	Adjusted R-squared:  0.005131
F-statistic: 1.239 on 36 and 1630 DF,  p-value: 0.1576
```

```
Call:
lm(formula = CAR5 ~ sentiment_risk_score + specificity_score +
    buzzwords_score_industry_norm + hedging_density_norma + candor_risk_component +
    log(1 + SIZE) + ROA + MOMENTUM + ESGscore + as.factor(Industry) +
    as.factor(year), data = merged_df)

Residuals:
      Min        1Q    Median        3Q       Max
-0.228440 -0.020274  0.000819  0.020273  0.223265

Coefficients:
                                                 Estimate Std. Error t value Pr(>|t|)
(Intercept)                                    -5.509e-02  2.338e-02  -2.357   0.0186 *
sentiment_risk_score                           -3.336e-03  1.492e-02  -0.224   0.8231
specificity_score                               2.861e-02  1.567e-02   1.825   0.0682 .
buzzwords_score_industry_norm                   2.875e-03  8.222e-03   0.350   0.7267
hedging_density_norma                           2.087e-02  1.119e-02   1.865   0.0624 .
candor_risk_component                           6.692e-03  1.459e-02   0.459   0.6466
log(1 + SIZE)                                   6.748e-04  8.350e-04   0.808   0.4191
ROA                                            -2.024e-04  1.756e-04  -1.153   0.2492
MOMENTUM                                       -3.292e-03  2.099e-03  -1.568   0.1170
ESGscore                                        2.698e-05  4.792e-04   0.056   0.9551
as.factor(Industry)Computers                   -4.988e-03  5.856e-03  -0.852   0.3944
as.factor(Industry)Durable manufacturers        2.692e-03  5.428e-03   0.496   0.6200
as.factor(Industry)Extractive industries        3.800e-03  7.334e-03   0.518   0.6045
as.factor(Industry)Financial institutions      -1.266e-02  6.222e-03  -2.035   0.0420 *
as.factor(Industry)Food                        -7.943e-03  6.773e-03  -1.173   0.2411
as.factor(Industry)Insurance and real estate   -1.000e-03  8.021e-03  -0.125   0.9008
as.factor(Industry)Mining and Construction     -1.122e-02  9.613e-03  -1.167   0.2435
as.factor(Industry)Pharmaceuticals              6.302e-04  6.620e-03   0.095   0.9242
as.factor(Industry)Retail                       8.759e-04  5.974e-03   0.147   0.8834
as.factor(Industry)Services                     1.301e-03  6.865e-03   0.190   0.8497
as.factor(Industry)Textiles, printing and publishing -3.151e-05  7.141e-03  -0.004   0.9965
as.factor(Industry)Transportation              -1.698e-03  6.529e-03  -0.260   0.7948
as.factor(Industry)Utilities                   -9.895e-03  7.325e-03  -1.351   0.1769
as.factor(year)2008                             2.608e-02  1.032e-02   2.527   0.0116 *
as.factor(year)2009                             1.877e-02  1.013e-02   1.853   0.0641 .
as.factor(year)2010                             1.878e-02  9.472e-03   1.982   0.0476 *
as.factor(year)2011                             2.244e-02  9.453e-03   2.374   0.0177 *
as.factor(year)2012                             1.655e-02  9.086e-03   1.822   0.0687 .
as.factor(year)2013                             2.060e-02  8.976e-03   2.296   0.0218 *
as.factor(year)2014                             1.862e-02  8.906e-03   2.091   0.0367 *
as.factor(year)2015                             1.774e-02  8.871e-03   2.000   0.0456 *
as.factor(year)2016                             1.901e-02  8.913e-03   2.133   0.0331 *
as.factor(year)2017                             1.344e-02  8.808e-03   1.526   0.1273
as.factor(year)2018                             1.544e-02  8.711e-03   1.773   0.0764 .
as.factor(year)2019                             1.363e-02  8.639e-03   1.578   0.1148
as.factor(year)2020                             1.268e-02  9.767e-03   1.298   0.1944
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.04089 on 1622 degrees of freedom
  (9 observations deleted due to missingness)
Multiple R-squared:  0.02756,	Adjusted R-squared:  0.00658
F-statistic: 1.314 on 35 and 1622 DF,  p-value: 0.1047
```

- CAR60 (GPT vs. NLP)

```
lm(formula = CAR60 ~ sentiment_score_gpt + specificity_score_gpt +
    buzzwords_score_gpt + hedging_score_gpt + candor_score_gpt +
    log(1 + SIZE) + ROA + MOMENTUM + ESGscore + as.factor(Industry) +
    as.factor(year), data = merged_df)

Residuals:
     Min       1Q   Median       3Q      Max
-0.55715 -0.05066  0.00073  0.04693  0.61810

Coefficients:
                                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                                 -0.0504498  0.0398405  -1.266    0.206
sentiment_score_gpt                          0.0473714  0.0700160   0.677    0.499
specificity_score_gpt                        0.0109221  0.0136970   0.797    0.425
buzzwords_score_gpt                          0.0139431  0.0387732   0.360    0.719
hedging_score_gpt                            0.0232595  0.0633871   0.367    0.714
candor_score_gpt                            -0.0219465  0.0250894  -0.875    0.382
log(1 + SIZE)                                0.0015880  0.0023811   0.667    0.505
ROA                                         -0.0001513  0.0004961  -0.305    0.760
MOMENTUM                                    -0.0351642  0.0059425  -5.917 3.98e-09 ***
ESGscore                                    -0.0003484  0.0013609  -0.256    0.798
as.factor(Industry)Computers                 0.0040382  0.0160261   0.252    0.801
as.factor(Industry)Durable manufacturers     0.0122646  0.0152481   0.804    0.421
as.factor(Industry)Extractive industries     0.0110901  0.0209026   0.531    0.596
as.factor(Industry)Financial institutions    0.0157891  0.0178338   0.885    0.376
as.factor(Industry)Food                      0.0108552  0.0179527   0.605    0.545
as.factor(Industry)Insurance and real estate 0.0158509  0.0215069   0.737    0.461
as.factor(Industry)Mining and Construction   0.0230001  0.0270992   0.849    0.396
as.factor(Industry)Others                   -0.0934838  0.0838711  -1.115    0.265
as.factor(Industry)Pharmaceuticals           0.0243350  0.0189003   1.288    0.198
as.factor(Industry)Retail                    0.0242966  0.0165298   1.470    0.142
as.factor(Industry)Services                  0.0108281  0.0193770   0.559    0.576
as.factor(Industry)Textiles, printing and publishing  0.0197612  0.0200546   0.985    0.325
as.factor(Industry)Transportation           -0.0079273  0.0185550  -0.427    0.669
as.factor(Industry)Utilities                 0.0016923  0.0194454   0.087    0.931
as.factor(year)2008                          0.0265398  0.0292619   0.907    0.365
as.factor(year)2009                          0.0089838  0.0287563   0.312    0.755
as.factor(year)2010                         -0.0026870  0.0268675  -0.100    0.920
as.factor(year)2011                          0.0020356  0.0268238   0.076    0.940
as.factor(year)2012                          0.0284177  0.0257790   1.102    0.270
as.factor(year)2013                         -0.0033047  0.0254795  -0.130    0.897
as.factor(year)2014                          0.0291004  0.0252809   1.151    0.250
as.factor(year)2015                          0.0230990  0.0251262   0.919    0.358
as.factor(year)2016                          0.0082935  0.0253099   0.328    0.743
as.factor(year)2017                          0.0023202  0.0249764   0.093    0.926
as.factor(year)2018                          0.0048145  0.0246998   0.195    0.845
as.factor(year)2019                          0.0099940  0.0245657   0.407    0.684
as.factor(year)2020                          0.0119633  0.0277937   0.430    0.667
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.116 on 1630 degrees of freedom
Multiple R-squared:  0.03785,	Adjusted R-squared:  0.0166
F-statistic: 1.781 on 36 and 1630 DF,  p-value: 0.003103
```

```
Call:
lm(formula = CAR60 ~ sentiment_risk_score + specificity_score +
    buzzwords_score_industry_norm + hedging_density_norma + candor_risk_component +
    log(1 + SIZE) + ROA + MOMENTUM + ESGscore + as.factor(Industry) +
    as.factor(year), data = merged_df)

Residuals:
     Min       1Q   Median       3Q      Max
-0.54145 -0.05055  0.00047  0.04816  0.60956

Coefficients:
                                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                                 -0.1093433  0.0662257  -1.651   0.0989 .
sentiment_risk_score                         0.0479624  0.0422784   1.134   0.2568
specificity_score                            0.0892314  0.0444061   2.009   0.0447 *
buzzwords_score_industry_norm               -0.0096849  0.0232931  -0.416   0.6776
hedging_density_norma                        0.0502529  0.0317085   1.585   0.1132
candor_risk_component                       -0.0165702  0.0413457  -0.401   0.6886
log(1 + SIZE)                                0.0018228  0.0023656   0.771   0.4411
ROA                                         -0.0001171  0.0004975  -0.235   0.8139
MOMENTUM                                    -0.0367402  0.0059468  -6.178 8.19e-10 ***
ESGscore                                    -0.0006732  0.0013576  -0.496   0.6201
as.factor(Industry)Computers                 0.0060317  0.0165902   0.364   0.7162
as.factor(Industry)Durable manufacturers     0.0132171  0.0153785   0.859   0.3902
as.factor(Industry)Extractive industries     0.0061491  0.0207789   0.296   0.7673
as.factor(Industry)Financial institutions    0.0135811  0.0176279   0.770   0.4412
as.factor(Industry)Food                      0.0136821  0.0191878   0.713   0.4759
as.factor(Industry)Insurance and real estate 0.0184727  0.0227232   0.813   0.4164
as.factor(Industry)Mining and Construction   0.0239973  0.0272353   0.881   0.3784
as.factor(Industry)Pharmaceuticals           0.0241717  0.0187543   1.289   0.1976
as.factor(Industry)Retail                    0.0227676  0.0169244   1.345   0.1787
as.factor(Industry)Services                  0.0045888  0.0194488   0.236   0.8135
as.factor(Industry)Textiles, printing and publishing  0.0212476  0.0202300   1.050   0.2937
as.factor(Industry)Transportation           -0.0114964  0.0184979  -0.622   0.5344
as.factor(Industry)Utilities                 0.0007975  0.0207510   0.038   0.9693
as.factor(year)2008                          0.0265736  0.0292463   0.909   0.3637
as.factor(year)2009                          0.0087010  0.0286953   0.303   0.7618
as.factor(year)2010                         -0.0036512  0.0268346  -0.136   0.8918
as.factor(year)2011                          0.0010346  0.0267798   0.039   0.9692
as.factor(year)2012                          0.0286551  0.0257398   1.113   0.2658
as.factor(year)2013                         -0.0039192  0.0254289  -0.154   0.8775
as.factor(year)2014                          0.0266139  0.0252312   1.055   0.2917
as.factor(year)2015                          0.0198484  0.0251311   0.790   0.4298
as.factor(year)2016                          0.0068348  0.0252511   0.271   0.7867
as.factor(year)2017                          0.0007896  0.0249530   0.032   0.9748
as.factor(year)2018                          0.0023059  0.0246779   0.093   0.9256
as.factor(year)2019                          0.0082902  0.0244750   0.339   0.7349
as.factor(year)2020                          0.0115646  0.0276695   0.418   0.6760
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1158 on 1622 degrees of freedom
  (9 observations deleted due to missingness)
Multiple R-squared:  0.03996,	Adjusted R-squared:  0.01925
F-statistic: 1.929 on 35 and 1622 DF,  p-value: 0.0009472
```

- **Appendix 12 - Model 4 Regressions (Section 4.5)**
    - CAR2 (GPT vs. NLP)

```
Call:
lm(formula = CAR2 ~ buzzwords_score_gpt * sentiment_score_gpt +
    specificity_score_gpt * ESGscore + hedging_score_gpt + candor_score_gpt +
    log(1 + SIZE) + ROA + MOMENTUM + as.factor(Industry) + as.factor(year),
    data = merged_df)

Residuals:
      Min        1Q    Median        3Q       Max
-0.183826 -0.012884 -0.000073  0.012518  0.211624

Coefficients:
                                               Estimate Std. Error t value Pr(>|t|)
(Intercept)                                  -1.358e-02  1.063e-02  -1.278  0.20159
buzzwords_score_gpt                           9.387e-03  9.864e-03   0.952  0.34140
sentiment_score_gpt                           1.709e-02  6.730e-02   0.254  0.79953
specificity_score_gpt                        -3.579e-03  8.982e-03  -0.398  0.69037
ESGscore                                     -9.392e-04  7.444e-04  -1.262  0.20728
hedging_score_gpt                             6.482e-03  1.692e-02   0.383  0.70167
candor_score_gpt                              1.093e-03  6.425e-03   0.170  0.86489
log(1 + SIZE)                                -4.973e-04  6.022e-04  -0.826  0.40905
ROA                                           5.592e-05  1.255e-04   0.446  0.65597
MOMENTUM                                     -3.261e-03  1.503e-03  -2.170  0.03017 *
as.factor(Industry)Computers                  4.978e-03  4.054e-03   1.228  0.21959
as.factor(Industry)Durable manufacturers      8.394e-03  3.859e-03   2.175  0.02975 *
as.factor(Industry)Extractive industries      5.935e-03  5.290e-03   1.122  0.26205
as.factor(Industry)Financial institutions    -4.298e-04  4.511e-03  -0.095  0.92412
as.factor(Industry)Food                      -1.585e-03  4.541e-03  -0.349  0.72706
as.factor(Industry)Insurance and real estate  2.917e-04  5.441e-03   0.054  0.95725
as.factor(Industry)Mining and Construction   -6.887e-03  6.855e-03  -1.005  0.31514
as.factor(Industry)Others                     4.514e-03  2.122e-02   0.213  0.83154
as.factor(Industry)Pharmaceuticals            5.942e-03  4.788e-03   1.241  0.21481
as.factor(Industry)Retail                     7.923e-03  4.181e-03   1.895  0.05826 .
as.factor(Industry)Services                   1.360e-03  4.903e-03   0.277  0.78157
as.factor(Industry)Textiles, printing and publishing  3.437e-03  5.073e-03   0.677  0.49825
as.factor(Industry)Transportation             3.148e-03  4.697e-03   0.670  0.50284
as.factor(Industry)Utilities                  2.731e-03  4.924e-03   0.555  0.57916
as.factor(year)2008                           2.124e-02  7.411e-03   2.866  0.00421 **
as.factor(year)2009                           1.615e-02  7.285e-03   2.217  0.02674 *
as.factor(year)2010                           1.237e-02  6.811e-03   1.816  0.06957 .
as.factor(year)2011                           1.740e-02  6.799e-03   2.559  0.01059 *
as.factor(year)2012                           1.026e-02  6.533e-03   1.571  0.11637
as.factor(year)2013                           1.398e-02  6.470e-03   2.161  0.03085 *
as.factor(year)2014                           1.239e-02  6.416e-03   1.931  0.05369 .
as.factor(year)2015                           1.526e-02  6.381e-03   2.392  0.01687 *
as.factor(year)2016                           1.261e-02  6.424e-03   1.963  0.04982 *
as.factor(year)2017                           1.055e-02  6.340e-03   1.664  0.09635 .
as.factor(year)2018                           1.255e-02  6.271e-03   2.001  0.04554 *
as.factor(year)2019                           9.843e-03  6.232e-03   1.579  0.11445
as.factor(year)2020                           1.276e-02  7.057e-03   1.808  0.07077 .
buzzwords_score_gpt:sentiment_score_gpt      -3.222e-02  9.371e-02  -0.344  0.73105
specificity_score_gpt:ESGscore                1.328e-03  1.508e-03   0.880  0.37876
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.02934 on 1628 degrees of freedom
Multiple R-squared:  0.03288,	Adjusted R-squared:  0.01031
F-statistic: 1.457 on 38 and 1628 DF,  p-value: 0.03637
```

```
Call:
lm(formula = CAR2 ~ buzzwords_score_industry_norm * sentiment_risk_score +
    specificity_score * ESGscore + hedging_density_norma + candor_risk_component +
    log(1 + SIZE) + ROA + MOMENTUM + as.factor(Industry) + as.factor(year),
    data = merged_df)

Residuals:
      Min        1Q    Median        3Q       Max
-0.184134 -0.013175 -0.000332  0.013136  0.208529

Coefficients:
                                                      Estimate Std. Error t value Pr(>|t|)
(Intercept)                                         -2.698e-02  2.207e-02  -1.223  0.22157
buzzwords_score_industry_norm                        1.859e-02  1.391e-02   1.337  0.18149
sentiment_risk_score                                 7.344e-03  2.515e-02   0.292  0.77033
specificity_score                                   -1.487e-02  2.596e-02  -0.573  0.56692
ESGscore                                            -4.097e-03  2.483e-03  -1.650  0.09912 .
hedging_density_norma                                1.617e-02  8.022e-03   2.015  0.04403 *
candor_risk_component                                1.222e-02  1.039e-02   1.176  0.23974
log(1 + SIZE)                                       -3.370e-04  5.946e-04  -0.567  0.57097
ROA                                                  4.768e-05  1.252e-04   0.381  0.70337
MOMENTUM                                            -3.493e-03  1.498e-03  -2.332  0.01982 *
as.factor(Industry)Computers                         2.968e-03  4.172e-03   0.711  0.47695
as.factor(Industry)Durable manufacturers             7.120e-03  3.874e-03   1.838  0.06627 .
as.factor(Industry)Extractive industries             5.321e-03  5.229e-03   1.018  0.30904
as.factor(Industry)Financial institutions           -1.555e-03  4.437e-03  -0.350  0.72606
as.factor(Industry)Food                             -4.159e-03  4.828e-03  -0.861  0.38911
as.factor(Industry)Insurance and real estate        -2.284e-03  5.711e-03  -0.400  0.68933
as.factor(Industry)Mining and Construction          -8.422e-03  6.854e-03  -1.229  0.21930
as.factor(Industry)Pharmaceuticals                   4.881e-03  4.714e-03   1.035  0.30065
as.factor(Industry)Retail                            8.647e-03  4.259e-03   2.030  0.04249 *
as.factor(Industry)Services                          4.384e-04  4.898e-03   0.089  0.92870
as.factor(Industry)Textiles, printing and publishing 1.594e-03  5.099e-03   0.313  0.75454
as.factor(Industry)Transportation                    1.628e-03  4.653e-03   0.350  0.72652
as.factor(Industry)Utilities                        -2.207e-03  5.223e-03  -0.423  0.67268
as.factor(year)2008                                  2.137e-02  7.360e-03   2.903  0.00375 **
as.factor(year)2009                                  1.652e-02  7.222e-03   2.287  0.02233 *
as.factor(year)2010                                  1.320e-02  6.748e-03   1.955  0.05071 .
as.factor(year)2011                                  1.827e-02  6.740e-03   2.710  0.00680 **
as.factor(year)2012                                  1.101e-02  6.478e-03   1.699  0.08942 .
as.factor(year)2013                                  1.438e-02  6.396e-03   2.249  0.02467 *
as.factor(year)2014                                  1.163e-02  6.347e-03   1.832  0.06716 .
as.factor(year)2015                                  1.549e-02  6.321e-03   2.450  0.01438 *
as.factor(year)2016                                  1.277e-02  6.352e-03   2.010  0.04460 *
as.factor(year)2017                                  1.098e-02  6.277e-03   1.749  0.08041 .
as.factor(year)2018                                  1.279e-02  6.206e-03   2.060  0.03955 *
as.factor(year)2019                                  9.711e-03  6.153e-03   1.578  0.11468
as.factor(year)2020                                  1.234e-02  6.956e-03   1.774  0.07625 .
buzzwords_score_industry_norm:sentiment_risk_score  -4.361e-02  4.529e-02  -0.963  0.33575
specificity_score:ESGscore                           6.581e-03  4.360e-03   1.509  0.13143
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.02911 on 1620 degrees of freedom
  (9 observations deleted due to missingness)
Multiple R-squared:  0.03897,	Adjusted R-squared:  0.01702
F-statistic: 1.775 on 37 and 1620 DF,  p-value: 0.00294
```

- CAR5 (GPT vs. NLP)

```
lm(formula = CAR5 ~ buzzwords_score_gpt * sentiment_score_gpt +
    specificity_score_gpt * ESGscore + hedging_score_gpt + candor_score_gpt +
    log(1 + SIZE) + ROA + MOMENTUM + as.factor(Industry) + as.factor(year),
    data = merged_df)

Residuals:
      Min        1Q    Median        3Q       Max
-0.228685 -0.020342  0.000284  0.020354  0.226344

Coefficients:
                                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                                 -2.181e-02  1.490e-02  -1.464   0.1435
buzzwords_score_gpt                          7.319e-03  1.383e-02   0.529   0.5967
sentiment_score_gpt                          2.147e-02  9.436e-02   0.228   0.8201
specificity_score_gpt                       -7.607e-03  1.259e-02  -0.604   0.5459
ESGscore                                    -9.340e-04  1.044e-03  -0.895   0.3710
hedging_score_gpt                            3.635e-02  2.372e-02   1.533   0.1256
candor_score_gpt                            -1.423e-03  9.008e-03  -0.158   0.8745
log(1 + SIZE)                                5.231e-04  8.443e-04   0.620   0.5356
ROA                                         -1.923e-04  1.760e-04  -1.093   0.2746
MOMENTUM                                    -3.102e-03  2.107e-03  -1.472   0.1413
as.factor(Industry)Computers                -3.807e-03  5.683e-03  -0.670   0.5031
as.factor(Industry)Durable manufacturers     2.941e-03  5.410e-03   0.544   0.5868
as.factor(Industry)Extractive industries     3.808e-03  7.416e-03   0.513   0.6077
as.factor(Industry)Financial institutions   -1.248e-02  6.325e-03  -1.973   0.0486 *
as.factor(Industry)Food                     -7.210e-03  6.367e-03  -1.132   0.2576
as.factor(Industry)Insurance and real estate -9.706e-05 7.628e-03  -0.013   0.9898
as.factor(Industry)Mining and Construction  -1.079e-02  9.610e-03  -1.122   0.2619
as.factor(Industry)Others                   -3.093e-02  2.975e-02  -1.040   0.2985
as.factor(Industry)Pharmaceuticals           8.432e-04  6.713e-03   0.126   0.9001
as.factor(Industry)Retail                    1.251e-04  5.861e-03   0.021   0.9830
as.factor(Industry)Services                  1.364e-03  6.874e-03   0.198   0.8427
as.factor(Industry)Textiles, printing and publishing  6.696e-04  7.113e-03   0.094   0.9250
as.factor(Industry)Transportation           -3.532e-04  6.586e-03  -0.054   0.9572
as.factor(Industry)Utilities                -5.861e-03  6.903e-03  -0.849   0.3960
as.factor(year)2008                          2.609e-02  1.039e-02   2.511   0.0121 *
as.factor(year)2009                          1.856e-02  1.021e-02   1.817   0.0693 .
as.factor(year)2010                          1.841e-02  9.550e-03   1.928   0.0540 .
as.factor(year)2011                          2.244e-02  9.532e-03   2.354   0.0187 *
as.factor(year)2012                          1.635e-02  9.160e-03   1.785   0.0745 .
as.factor(year)2013                          2.094e-02  9.072e-03   2.309   0.0211 *
as.factor(year)2014                          2.053e-02  8.995e-03   2.282   0.0226 *
as.factor(year)2015                          1.812e-02  8.946e-03   2.026   0.0430 *
as.factor(year)2016                          1.908e-02  9.007e-03   2.118   0.0343 *
as.factor(year)2017                          1.362e-02  8.888e-03   1.532   0.1257
as.factor(year)2018                          1.575e-02  8.792e-03   1.791   0.0735 .
as.factor(year)2019                          1.404e-02  8.738e-03   1.607   0.1083
as.factor(year)2020                          1.331e-02  9.894e-03   1.345   0.1787
buzzwords_score_gpt:sentiment_score_gpt     -5.887e-02  1.314e-01  -0.448   0.6542
specificity_score_gpt:ESGscore               2.200e-03  2.115e-03   1.040   0.2984
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.04113 on 1628 degrees of freedom
Multiple R-squared:  0.02744,   Adjusted R-squared:  0.004737
F-statistic: 1.209 on 38 and 1628 DF,  p-value: 0.1802
```

```
Call:
lm(formula = CAR5 ~ buzzwords_score_industry_norm * sentiment_risk_score +
    specificity_score * ESGscore + hedging_density_norma + candor_risk_component +
    log(1 + SIZE) + ROA + MOMENTUM + as.factor(Industry) + as.factor(year),
    data = merged_df)

Residuals:
      Min        1Q    Median        3Q       Max
-0.228274 -0.020145  0.000862  0.020704  0.224590

Coefficients:
                                                      Estimate Std. Error t value Pr(>|t|)
(Intercept)                                         -0.0242650  0.0309753  -0.783  0.43353
buzzwords_score_industry_norm                        0.0077334  0.0195232   0.396  0.69208
sentiment_risk_score                                 0.0046871  0.0353060   0.133  0.89440
specificity_score                                   -0.0317299  0.0364439  -0.871  0.38407
ESGscore                                            -0.0063531  0.0034852  -1.823  0.06851 .
hedging_density_norma                                0.0214254  0.0112616   1.903  0.05728 .
candor_risk_component                                0.0070635  0.0145884   0.484  0.62832
log(1 + SIZE)                                        0.0007043  0.0008347   0.844  0.39892
ROA                                                 -0.0002127  0.0001757  -1.210  0.22632
MOMENTUM                                            -0.0035158  0.0021024  -1.672  0.09467 .
as.factor(Industry)Computers                        -0.0051281  0.0058564  -0.876  0.38135
as.factor(Industry)Durable manufacturers             0.0020137  0.0054379   0.370  0.71120
as.factor(Industry)Extractive industries             0.0030888  0.0073408   0.421  0.67398
as.factor(Industry)Financial institutions           -0.0132614  0.0062281  -2.129  0.03338 *
as.factor(Industry)Food                             -0.0085099  0.0067771  -1.256  0.20941
as.factor(Industry)Insurance and real estate        -0.0010867  0.0080172  -0.136  0.89219
as.factor(Industry)Mining and Construction          -0.0119018  0.0096211  -1.237  0.21625
as.factor(Industry)Pharmaceuticals                   0.0004592  0.0066173   0.069  0.94468
as.factor(Industry)Retail                            0.0003600  0.0059786   0.060  0.95199
as.factor(Industry)Services                          0.0006537  0.0068755   0.095  0.92426
as.factor(Industry)Textiles, printing and publishing -0.0010267  0.0071572  -0.143  0.88595
as.factor(Industry)Transportation                   -0.0020986  0.0065317  -0.321  0.74803
as.factor(Industry)Utilities                        -0.0106324  0.0073325  -1.450  0.14724
as.factor(year)2008                                  0.0266505  0.0103312   2.580  0.00998 **
as.factor(year)2009                                  0.0194336  0.0101385   1.917  0.05544 .
as.factor(year)2010                                  0.0193242  0.0094730   2.040  0.04152 *
as.factor(year)2011                                  0.0233395  0.0094610   2.467  0.01373 *
as.factor(year)2012                                  0.0174630  0.0090940   1.920  0.05500 .
as.factor(year)2013                                  0.0212615  0.0089789   2.368  0.01800 *
as.factor(year)2014                                  0.0192868  0.0089094   2.165  0.03055 *
as.factor(year)2015                                  0.0183298  0.0088726   2.066  0.03900 *
as.factor(year)2016                                  0.0196386  0.0089162   2.203  0.02776 *
as.factor(year)2017                                  0.0140946  0.0088121   1.599  0.10991
as.factor(year)2018                                  0.0159890  0.0087121   1.835  0.06665 .
as.factor(year)2019                                  0.0139424  0.0086371   1.614  0.10667
as.factor(year)2020                                  0.0130912  0.0097648   1.341  0.18022
buzzwords_score_industry_norm:sentiment_risk_score  -0.0184133  0.0635803  -0.290  0.77215
specificity_score:ESGscore                           0.0113087  0.0061210   1.848  0.06485 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.04087 on 1620 degrees of freedom
  (9 observations deleted due to missingness)
Multiple R-squared:  0.0297,    Adjusted R-squared:  0.007541
F-statistic:  1.34 on 37 and 1620 DF,  p-value: 0.08395
```

- CAR60 (GPT vs. NLP)

```
lm(formula = CAR60 ~ buzzwords_score_gpt * sentiment_score_gpt +
    specificity_score_gpt * ESGscore + hedging_score_gpt + candor_score_gpt +
    log(1 + SIZE) + ROA + MOMENTUM + as.factor(Industry) + as.factor(year),
    data = merged_df)

Residuals:
     Min       1Q   Median       3Q      Max
-0.55864 -0.05030  0.00061  0.04722  0.61610

Coefficients:
                                                      Estimate Std. Error t value Pr(>|t|)
(Intercept)                                         -0.0410683  0.0420287  -0.977    0.329
buzzwords_score_gpt                                  0.0165593  0.0390139   0.424    0.671
sentiment_score_gpt                                  0.1625931  0.2661847   0.611    0.541
specificity_score_gpt                               -0.0199441  0.0355268  -0.561    0.575
ESGscore                                            -0.0028897  0.0029444  -0.981    0.327
hedging_score_gpt                                    0.0357584  0.0669176   0.534    0.593
candor_score_gpt                                    -0.0201452  0.0254130  -0.793    0.428
log(1 + SIZE)                                        0.0016069  0.0023818   0.675    0.500
ROA                                                 -0.0001516  0.0004964  -0.305    0.760
MOMENTUM                                            -0.0352668  0.0059448  -5.932 3.64e-09 ***
as.factor(Industry)Computers                         0.0040609  0.0160332   0.253    0.800
as.factor(Industry)Durable manufacturers             0.0117715  0.0152621   0.771    0.441
as.factor(Industry)Extractive industries             0.0102696  0.0209221   0.491    0.624
as.factor(Industry)Financial institutions            0.0153771  0.0178430   0.862    0.389
as.factor(Industry)Food                              0.0105232  0.0179615   0.586    0.558
as.factor(Industry)Insurance and real estate         0.0158667  0.0215194   0.737    0.461
as.factor(Industry)Mining and Construction           0.0225332  0.0271108   0.831    0.406
as.factor(Industry)Others                           -0.0910921  0.0839199  -1.085    0.278
as.factor(Industry)Pharmaceuticals                   0.0248508  0.0189381   1.312    0.190
as.factor(Industry)Retail                            0.0244493  0.0165348   1.479    0.139
as.factor(Industry)Services                          0.0103666  0.0193921   0.535    0.593
as.factor(Industry)Textiles, printing and publishing  0.0196344  0.0200650   0.979    0.328
as.factor(Industry)Transportation                   -0.0077012  0.0185785  -0.415    0.679
as.factor(Industry)Utilities                         0.0013734  0.0194741   0.071    0.944
as.factor(year)2008                                  0.0274231  0.0293136   0.936    0.350
as.factor(year)2009                                  0.0096850  0.0288122   0.336    0.737
as.factor(year)2010                                 -0.0029446  0.0269404  -0.109    0.913
as.factor(year)2011                                  0.0025863  0.0268913   0.096    0.923
as.factor(year)2012                                  0.0292988  0.0258406   1.134    0.257
as.factor(year)2013                                 -0.0026582  0.0255919  -0.104    0.917
as.factor(year)2014                                  0.0295737  0.0253754   1.165    0.244
as.factor(year)2015                                  0.0240734  0.0252364   0.954    0.340
as.factor(year)2016                                  0.0087526  0.0254101   0.344    0.731
as.factor(year)2017                                  0.0027500  0.0250741   0.110    0.913
as.factor(year)2018                                  0.0050317  0.0248031   0.203    0.839
as.factor(year)2019                                  0.0101952  0.0246506   0.414    0.679
as.factor(year)2020                                  0.0126986  0.0279107   0.455    0.649
buzzwords_score_gpt:sentiment_score_gpt             -0.1623827  0.3706406  -0.438    0.661
specificity_score_gpt:ESGscore                       0.0058353  0.0059658   0.978    0.328
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.116 on 1628 degrees of freedom
Multiple R-squared:  0.03856,   Adjusted R-squared:  0.01612
F-statistic: 1.718 on 38 and 1628 DF,  p-value: 0.004394
```

```
Call:
lm(formula = CAR60 ~ buzzwords_score_industry_norm * sentiment_risk_score +
    specificity_score * ESGscore + hedging_density_norma + candor_risk_component +
    log(1 + SIZE) + ROA + MOMENTUM + as.factor(Industry) + as.factor(year),
    data = merged_df)

Residuals:
     Min       1Q   Median       3Q      Max
-0.54311 -0.05138  0.00068  0.04918  0.61187

Coefficients:
                                                      Estimate Std. Error t value Pr(>|t|)
(Intercept)                                         -0.0902580  0.0877977  -1.028    0.304
buzzwords_score_industry_norm                        0.0425495  0.0553375   0.769    0.442
sentiment_risk_score                                 0.1412678  0.1000730   1.412    0.158
specificity_score                                    0.0117619  0.1032981   0.114    0.909
ESGscore                                            -0.0091709  0.0098786  -0.928    0.353
hedging_density_norma                                0.0479613  0.0319203   1.503    0.133
candor_risk_component                               -0.0160735  0.0413501  -0.389    0.698
log(1 + SIZE)                                        0.0018701  0.0023660   0.790    0.429
ROA                                                 -0.0001522  0.0004981  -0.306    0.760
MOMENTUM                                            -0.0372661  0.0059592  -6.254 5.12e-10 ***
as.factor(Industry)Computers                         0.0053485  0.0165996   0.322    0.747
as.factor(Industry)Durable manufacturers             0.0119493  0.0154136   0.775    0.438
as.factor(Industry)Extractive industries             0.0051491  0.0208072   0.247    0.805
as.factor(Industry)Financial institutions            0.0123744  0.0176531   0.701    0.483
as.factor(Industry)Food                              0.0124889  0.0192093   0.650    0.516
as.factor(Industry)Insurance and real estate         0.0182234  0.0227242   0.802    0.423
as.factor(Industry)Mining and Construction           0.0237086  0.0272704   0.869    0.385
as.factor(Industry)Pharmaceuticals                   0.0239564  0.0187564   1.277    0.202
as.factor(Industry)Retail                            0.0222473  0.0169460   1.313    0.189
as.factor(Industry)Services                          0.0042235  0.0194882   0.217    0.828
as.factor(Industry)Textiles, printing and publishing  0.0198524  0.0202867   0.979    0.328
as.factor(Industry)Transportation                   -0.0125548  0.0185136  -0.678    0.498
as.factor(Industry)Utilities                        -0.0006988  0.0207835  -0.034    0.973
as.factor(year)2008                                  0.0285516  0.0292831   0.975    0.330
as.factor(year)2009                                  0.0087293  0.0287370   0.304    0.761
as.factor(year)2010                                 -0.0031354  0.0268506  -0.117    0.907
as.factor(year)2011                                  0.0021376  0.0268167   0.080    0.936
as.factor(year)2012                                  0.0300684  0.0257764   1.167    0.244
as.factor(year)2013                                 -0.0031219  0.0254501  -0.123    0.902
as.factor(year)2014                                  0.0273923  0.0252532   1.085    0.278
as.factor(year)2015                                  0.0205163  0.0251488   0.816    0.415
as.factor(year)2016                                  0.0074802  0.0252724   0.296    0.767
as.factor(year)2017                                  0.0014201  0.0249773   0.057    0.955
as.factor(year)2018                                  0.0028947  0.0246939   0.117    0.907
as.factor(year)2019                                  0.0085245  0.0244815   0.348    0.728
as.factor(year)2020                                  0.0120221  0.0276778   0.434    0.664
buzzwords_score_industry_norm:sentiment_risk_score  -0.1889036  0.1802147  -1.048    0.295
specificity_score:ESGscore                           0.0150373  0.0173496   0.867    0.386
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1158 on 1620 degrees of freedom
  (9 observations deleted due to missingness)
Multiple R-squared:  0.04113,   Adjusted R-squared:  0.01922
F-statistic: 1.878 on 37 and 1620 DF,  p-value: 0.001163
```

- **Appendix 13 - R Code used for Greenwashing Detection and Market Reaction Analysis**

(see attached annex)

This annex provides the full R script used to compute the linguistic indicators (sentiment, specificity, buzzword density, hedging, and candor), construct the rule-based Greenwashing Score, and perform the regression analyses linking these scores to stock market reactions (CAR).