The Truth behind ESG Disclosures: Detecting Greenwashing through Text Readability

Sitian Li *

International Business college, Dongbei University of Finance and Economics, 116025 No. 217, Jianshan Street, Shahekou District, Dalian City, Liaoning Province, China

* Corresponding Author Email: Sitianli_lst@outlook.com

Abstract. As the concept of sustainable development advances, ESG disclosure has become a key component of corporate non-financial reporting. However, some firms engage in strategic obfuscation by complicating the language of ESG reports to exaggerate their actual performance, resulting in greenwashing. This study examines 1133 non-financial A-share listed firms in China from 2019 to 2023, extracting ESG sections from annual reports to construct a text readability indicator. Empirical findings suggest that lower readability is significantly associated with a higher degree of greenwashing. This effect is mitigated in larger firms, where reputation and regulatory pressure impose constraints. Further heterogeneity analysis reveals that such textual manipulation is more prevalent in high-pollution sectors. The study sheds light on the "soft manipulation" path of ESG disclosures, enriching the literature on greenwashing and offering language-based warning indicators for regulators.

Keywords: ESG Disclosures, Greenwashing, Text Readability, non-financial A-share listed firms.

1. Introduction

Against the backdrop of accelerating global sustainability strategies, ESG (Environmental, Social, Governance) disclosure has emerged as a core element in evaluating corporate non-financial performance. In April 2024, the China Securities Regulatory Commission (CSRC) released the Guidelines for Listed Companies' Sustainable Development Reporting (Trial), signalling a shift toward the institutionalization and standardization of ESG practices in China.

However, in the absence of uniform disclosure standards and mandatory frameworks, some firms engage in "greenwashing" behaviors, manipulating language in ESG reports to maintain a favourable public image or to gain policy benefits, rather than genuinely improving ESG performance.

This study is expected to focus on the textual readability of ESG disclosures and proposes a framework for identifying "linguistic greenwashing". Using a panel regression model, it empirically examines the impact of readability on greenwashing behavior and explores the moderating role of firm characteristics, such as firm size.

2. Theoretical analysis and research hypotheses

2.1. Theoretical foundations

Corporate information disclosure behavior is shaped by multiple theoretical perspectives, primarily including signalling theory, agency theory, and institutional theory. Signalling theory emphasizes that firms convey their intrinsic value and quality to external stakeholders by releasing strategic information. Agency theory focuses on information asymmetry and the resulting problems between principals and agents [1]. Institutional theory suggests that organizational behavior is influenced by external norms, regulations, and the broader cultural and legal environment.

In the context of ESG disclosure in China, where reporting requirements are not fully mandatory and evaluation standards remain inconsistent, managers who possess an informational advantage may adjust textual structures and linguistic styles to influence external perceptions. This form of soft manipulation based on linguistic complexity has become a low-cost means for firms to capitalize on

information asymmetry. These findings indicate that accurate and transparent ESG reporting provides a useful guidance for enhancing corporate value.

2.2. Literature review and comparative analysis

Existing literature has largely focused on rating quality, completeness of disclosure, and its impact on capital markets, particularly highlighting how ESG ratings can influence firm valuation, reduce finance costs, and serve as effective signals to external stakeholders [2,3]. While several scholars have laid the groundwork for readability research, relatively few studies have yet to be systematically applied to Chinese ESG reports [4,5]. As a result, there is a lack of empirical evidence on the relationship between textual complexity and greenwashing behavior.

Existing studies suggest that higher readability is associated with better information efficiency [6]. In general, firms with strong growth prospects are more likely to publish transparent and readable reports in order to enhance external recognition, build reputation, and secure business advantages [7]. However, attributes such as firm size and profitability may moderate the extent of strategic linguistic manipulation, yet quantitative analysis in this area remains scarce. These gaps highlight the need for further empirical investigation into how textual complexity may function as a soft mechanism of impression management, particularly in emerging markets.

2.3. Research hypothesis

The study constructs a readability index based on the original Chinese text of annual ESG reoorts and incorporates key firm-level variables to address the lack of heterogeneity analysis in existing research on linguistic greenwashing. Accordingly, the following hypothesis is proposed:

H1. All else being equal, firms with lower ESG report readability are more likely to engage in greenwashing behavior.

3. Methodology

3.1. Sample and variable description

This sample consists of non-financial firms listed on the Shanghai and Shenzhen A-share markets from 2019 to 2023. The final sample includes 1133 firms and 5664 panel observations. The main data sources include Wind database (used for financial variables and preprocessing of readability metrics), CSMAR (for the construction of the readability index), and full-text corporate annual reports.

To ensure data quality and consistency, the following filtering criteria were applied:

(1) Firms subject to special treatment (ST or ST*) during the sample period were excluded;

(2) Financial institutions such as banks and insurance companies were removed due to their distinct reporting structure.

(3) Firms with missing key variables were excluded from the final dataset.

3.2. Variable definition and measurement

This study constructs a panel regression model to examine the relationship between ESG disclosure readability and greenwashing behavior, while considering firm size as a potential moderating factor. Based on prior forward-looking disclosure research, the definitions and measurement approaches of the variables are summarized in Table 1 [8].

To enhance robustness, variables such as ROE and Ret_Vol were minorized at the 1st and 99th percentiles. Pearson correlation coefficients among variables are all below 0.4, and variance inflation factors (VIFs) are below 2, indicating no significant multicollinearity.

Table 1. Variable Definitions and Measurement

Type	Variable Symbol	Decription
Dependent Variable	GW_Score	The difference between the Wind ESG overall score and the ESG management score, used to measure the degree of ESG "greenwashing" or superficial packaging
Main Independent Variable	RA	Gunning Fog Index of the ESG-related paragraphs in annual reports; higher values indicate lower readability (a negative readability indicator)
Moderating Variable	RA × Size	Interaction term used to test whether firm size moderates the relationship between readability and greenwashing behavior
Control Variable	Size	Natural logarithm of total assets
Control Variable	Lev	Total liabilities divided by total assets, capturing solvency risk
Control Variable	ROE	Return on equity, representing firm profitability
Control Variable	Ret_Vol	Standard deviation of annual stock returns, measuring market volatility risk
Fixed Effect	γ_t	Year fixed effects to control for systematic policy or macroeconomic changes
Fixed Effect	μ_i	Firm fixed effects to account for unobservable firm-specific characteristics

4. Empirical results and analysis

4.1. Model specification

To empirically test the proposed hypothesis, the following panel regression model with firm and year fixed effects is constructed:

$$\text{GW\_Score}_{it} = \beta_0 + \beta_1 RA_{it} + \beta_2 Size_{it} + \beta_3 (RA \times Size)_{it} + \beta_4 X_{it} + \mu_i + \gamma_t + \varepsilon_{it} \quad (1)$$

Where X_{it} includes control variables.

4.2. Main regression results

Table 2 reports the estimation results from the fixed effects (FE), random effects (RE), and interaction regression models. The primary focus is to examine the impact of ESG disclosure readability (RA) on greenwashing behavior (GW_Score), along with the moderating effect of firm size.

The regression results indicate that textual readability has a statistically significant and positive association with greenwashing behavior, supporting the hypothesis that complex language increases strategic disclosure [3]. Moreover, the interaction term RA × Size shows a significantly negative effect, suggesting that larger firms are less likely to rely on language manipulation to engage in greenwashing. This outcome implies that firms under greater institutional pressure and public scrutiny face higher reputation risks, which in turn reduces the space for linguistic opportunism.

Among the control variables, leverage and market volatility consistently constrain greenwashing tendencies, whereas profitability (ROE) appears to enhance firms' capacity to manage impressions through ESG narratives.

These findings remain robust after controlling for year and firm fixed effects, effectively mitigating the influence of macroeconomic trends and industry-specific fluctuations.

Table 2. Results of Main Regression and Interaction Effects

Variables	FE Model	RE Model	Interaction Model
RA	0.00186**	0.00238***	0.00282***
	(0.00082)	(0.00073)	(0.00074)
RA × Size	---	---	-0.000809***
			(0.00023)
Size	-0.0151***	-0.0184***	-0.0236***
	(0.00205)	(0.00165)	(0.00225)
Lev	-0.000546***	-0.000490***	-0.000505***
	(0.00014)	(0.00012)	(0.00012)
ROA	0.0000986	0.000266	0.000176
	(0.00053)	(0.00046)	(0.00047)
Ret_Vol	-0.000365**	-0.000534***	-0.000533***
	(0.00015)	(0.00013)	(0.00013)
ROE	0.000582**	0.000614**	0.000636***
	(0.00027)	(0.00024)	(0.00024)
Year fixed effects	YES	YES	YES
Observations	5664	5664	5664
Firms	1133	1133	1133

4.3. Descriptive statistics

Table 3 presents the year-by-year descriptive statistics of the main variables from 2019 to 2023, while Table 4 provides a summary across the full sample. The average GW_Score remains relatively stable between 2.88 and 2.92, suggesting a consistent pattern in green disclosure behaviors across firms. However, the RA index shows a marked decline in 2023, implying increased textual complexity in ESG reporting and a potential trend toward obfuscatory disclosure.

In terms of financial characteristics, both ROA and ROE exhibit large standard deviations and wide value ranges, reflecting significant disparities in profitability and exposure to operational risk. The average leverage ratio (Lev) has sown a modest upward trend, suggesting rising debt pressure for some firms.

Overall, the data demonstrate a pattern of convergence in outward ESG disclosure and divergence in underlying financial performance. This observation aligns with prior studies, and supports the empirical foundation for examing potential greenwashing behaviors and disclosure strategies [9,10].

Table 3. Descriptive Statistics by Year

	(1)	(2)	(3)	(4)	(5)
Variables	n	mean	sd	min	max
GW_Score2023	1,133.00	2.92	0.11	1.90	3.01
RA2023	1,133.00	-32.06	3.11	-44.90	-23.00
size2023	1,133.00	22.76	1.42	19.47	28.34
Lev2023	1,133.00	43.83	19.09	2.11	120.37
ROA2023	1,133.00	3.56	6.57	-28.63	49.41
Ret_Vol2023	1,133.00	34.79	12.16	12.01	83.21
ROE2023	1,133.00	4.48	12.66	-145.46	67.05
GW_Score2022	1,133.00	2.88	0.13	1.18	3.01
RA2022	1,133.00	-20.02	1.88	-29.21	-14.71
size2022	1,133.00	22.48	1.25	19.54	28.30
Lev2022	1,133.00	41.58	17.48	5.50	98.08
ROA2022	1,133.00	4.99	8.32	-29.70	70.33
Ret_Vol2022	1,133.00	46.71	11.15	18.82	105.67
ROE2022	1,133.00	6.19	15.22	-149.92	80.65
GW_Score2021	1,133.00	2.88	0.14	1.47	3.01
RA2021	1,133.00	-20.58	1.98	-38.31	-16.03
size2021	1,133.00	22.39	1.24	19.61	28.27
Lev2021	1,133.00	40.90	17.38	3.17	92.87
ROA2021	1,133.00	5.96	8.15	-76.41	52.73
Ret_Vol2021	1,133.00	47.89	12.74	15.46	93.24
ROE2021	1,133.00	7.92	14.91	-139.95	151.77
GW_Score2020	1,133.00	2.92	0.13	1.46	3.01
RA2020	1,133.00	-19.78	2.00	-35.32	-15.26
size2020	1,133.00	22.26	1.22	19.81	28.18
Lev2020	1,133.00	39.52	17.38	2.69	97.89
ROA2020	1,133.00	5.98	8.11	-64.39	81.18
Ret_Vol2020	1,133.00	49.02	10.98	17.31	94.54
ROE2020	1,133.00	7.74	15.46	-191.47	238.62
GW_Score2019	1,133.00	2.90	0.16	1.24	3.01
RA2019	1,133.00	-19.77	2.98	-37.69	-14.39
size2019	1,133.00	22.40	1.35	19.53	28.34
Lev2019	1,133.00	41.04	18.97	3.99	97.63
ROA2019	1,133.00	5.88	7.60	-64.49	44.66
Ret_Vol2019	1,133.00	42.73	19.40	14.66	405.83
ROE2019	1,133.00	7.27	13.42	-169.87	70.35

Table 4. Overall Summary Statistics of Variables

Variables	mean	(sd)
GW_Score	2.901	(0.136)
RA	-22.44	(5.403)
RA_size_interaction	-11.61	(12.29)
size	22.46	(1.310)
Lev	41.36	(18.09)
ROA	5.281	(7.820)
Ret_Vol	44.23	(14.59)
ROE	6.721	(14.42)
Observations	5,664	5,664
Number of firm_num	1,133	1,133

4.4. Industry-level heterogeneity analysis

This study further investigates industry-level heterogeneity by dividing the sample into 12 industry groups based on the China Securities Regulatory Commission (CSRC) classification. Fixed-effects regression models were estimated separately for each group, and the coefficients and significance levels of key variables are reported in Table 5.

Table 5. Regression Results by Industry

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Variables	Industry A	Industry B	Industry C	Industry D	Industry E	Industry F	Industry G	Industry I	Industry K	Industry L
RA	-0.002	-0.001	0.001	-0.001	0.006	0.002	-0.000	-0.002	-0.001	-0.027**
	(0.005)	(0.004)	(0.001)	(0.003)	(0.005)	(0.003)	(0.003)	(0.002)	(0.002)	(0.010)
size	-0.011	-0.039***	-0.012***	-0.005	0.000	-0.017*	-0.007	-0.006	-0.013	0.004
	(0.010)	(0.012)	(0.002)	(0.006)	(0.010)	(0.009)	(0.008)	(0.006)	(0.008)	(0.021)
Lev	-0.002***	0.001*	-0.000***	-0.000	-0.001*	-0.000	-0.002***	-0.001	-0.000	0.000
	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)	(0.002)
ROA	-0.004	0.003	0.000	0.000	-0.001	0.001	-0.005***	-0.000	-0.005	0.011
	(0.004)	(0.004)	(0.000)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.012)
Ret_Vol	-0.000	-0.000	-0.001***	-0.001	0.000	-0.001	-0.001	-0.000	0.000	-0.005**
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
ROE	0.001	-0.001	0.001**	0.000	0.000	-0.000	0.002*	0.001	0.003*	-0.003
	(0.002)	(0.002)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.006)
2020.year	-0.058**	-0.082*	-0.036***	0.074***	0.047	0.082***	0.107***	0.255***	0.587***	1.273***
	(0.029)	(0.043)	(0.005)	(0.019)	(0.030)	(0.026)	(0.024)	(0.019)	(0.024)	(0.102)
2021.year	-0.108***	-0.143***	-0.076***	0.029	0.058*	0.046*	0.085***	0.220***	0.535***	1.257***
	(0.027)	(0.043)	(0.005)	(0.019)	(0.031)	(0.025)	(0.024)	(0.019)	(0.025)	(0.098)
2022.year	-0.083***	-0.143***	-0.068***	-0.013	0.045	0.002	0.068***	0.198***	0.525***	1.266***
	(0.026)	(0.043)	(0.005)	(0.019)	(0.031)	(0.025)	(0.025)	(0.019)	(0.024)	(0.101)
2023.year	-0.116	-0.079	-0.022**	0.052	0.125*	0.103**	0.125***	0.233***	0.557***	0.982***
	(0.074)	(0.066)	(0.011)	(0.041)	(0.071)	(0.047)	(0.042)	(0.034)	(0.034)	(0.107)
Constant	3.313***	3.792***	3.287***	3.014***	3.006***	3.274***	3.067***	2.784***	2.660***	1.228*
	(0.214)	(0.255)	(0.038)	(0.142)	(0.224)	(0.187)	(0.184)	(0.143)	(0.193)	(0.653)
Observations	90	155	3,950	200	175	260	204	465	110	45
Number of firms	18	31	790	40	35	52	41	93	22	9

Standard errors in parentheses

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

The coefficient of RA shows a positive relationship with greenwashing behavior in several industries. In particular, the manufacturing sector (Industry C) and the leasing and business service (Industry L) exhibit significantly positive coefficients, indicating a stronger tendency among firms in these sectors to used linguistic complexity as a form of strategic impression management in ESG disclosures. These findings align with the characteristics pressure of manufacturing sector, where firms face high external ESG expectations but encounter substantial costs in achieving substantive improvements [11].

By contrast, for the Accommodation and Catering Services industry (H), the small sample size (n = 10) limits the readability of regression estimates. As a result, the findings are excluded from further analysis and interpretation.

Firm size exhibits a generally negative relationship with greenwashing across most industries, indicating that larger firms tend to show weaker greenwashing tendencies, possibly due to higher reputational risk and stronger institutional constraints. Control variables such as ROE and Ret_Vol remain directionally consistent across industries, lending further robustness to the empirical findings.

4.5. Firm size heterogeneity analysis

To further investigate whether firm size moderates' firms' reliance on textual complexity in ESG disclosure, the sample is divided into small and large firms based on the median of total assets, and conduct random-effects regressions for each subgroup.

As shown in the Table 6, the coefficient of RA is significantly higher for small firms (0.00293, $p < 0.05$) compared to large firms (0.00191, $p < 0.1$), indicating a strong association between low

readability and greenwashing behavior among small firms. This may suggest that SMEs, facing greater resource constraints, are more inclined to adopt linguistic complexity as a cost-effective strategy to enhance the perceived quality of ESG disclosures.

Interestingly, the coefficient on ROA is positively and significant in the small-firm group, but negative in the large-firm group, implying that profitability may incentivize strategic disclosures differently depending on firm scale. Similarly, ROE is only significant in the large-firm sample, indicating that larger, more established firms may have stronger reputational motivations for ESG transparency.

Table 6. Regression Results by Firm Size Group

Variables	Small Firms (n=2,831)	Large Firms (n=2,833)
RA	0.00293** (0.00104)	0.00191* (0.00102)
Size	-0.00577 (0.00451)	-0.02690*** (0.00283)
Lev	-0.00057*** (0.00015)	-0.00056*** (0.00019)
ROA	0.00152** (0.00055)	-0.00199** (0.00079)
Ret_Vol	-0.00073*** (0.00018)	-0.00037* (0.00020)
ROE	0.00000 (0.00028)	0.00164*** (0.00042)
2020.year	0.01577** (0.00663)	0.02729*** (0.00903)
2021.year	-0.01886*** (0.00684)	-0.01051 (0.00875)
2022.year	-0.01681** (0.00692)	-0.01446* (0.00852)
2023.year	0.04510*** (0.01484)	0.06941*** (0.01477)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

5. Discussion

5.1. Casual interpretation

Despite growing regulatory emphasis on ESG standardization in recent years, firms may continue to engage in greenwashing by manipulating textual complexity to obscure their actual performance in environmental, social, and governance (ESG) domains. Anchored in Readability Theory and lens of strategic information manipulation, this study identifies textual complexity as a covert pathway for greenwashing. In essence, firms may substitute linguistic sophistication for substantive ESG improvements as a strategic response to external scrutiny and internal constraints.

The main regression results show a significant positive association between ESG disclosure complexity (measured by RA) and the likelihood of greenwashing (proxied by GW_Score). This finding suggests that less readable ESG reports are more likely to conceal poor sustainability performance or selectively highlight favorable information. This supports the "language filtering effect" posited by Strategic Disclosure Theory, whereby firms deliberately raise the cognitive threshold for external stakeholders to critically assess underlying performance [11].

Further analysis incorporating firm size interaction terms reveals that this association weakens significantly for large firms. This can be interpreted from two perspectives:

(1) Reputational accountability: large firms are more visible and subject to media and investor scrutiny, making inconsistency between claims and actions more costly.

(2) Compliance burden: large firms are more likely to pursue "substantive compliance" over "symbolic disclosure" due to more rigorous reporting standards and internal governance mechanisms. These insights complement the Legitimacy Theory, particularly its notion of context-contingent strategic responses.

5.2. Sector and firm-level heterogeneity

Industry-specific regression results reveal pronounced variation in the relationship between disclosure readability and greenwashing behaviors. Notably, firms in the manufacturing sector which face both high-emission and high- scrutiny, show a significantly positive RA coefficient. This

suggests that they are more likely to rely on complex ESG narratives to dilute stakeholder perceptions of negative impacts. This aligns with the logic of Expectation-Performance Disclosure Management, wherein firms manage stakeholder expectations through symbolic communication rather than substantive improvements [12]. In contrast, firms in the Accommodation and Catering Services sector do not exhibit significant textual manipulation effects, potentially due to stronger regulatory boundaries and policy-driven disclosure frameworks that limit the degree of narrative discretion.

Firm size also plays a decisive role. Regression results from sub-sample analysis indicate that small and medium-sized enterprises (SMEs) exhibit a higher sensitivity to textual manipulation in ESG reform and greater incentive to symbolically comply with disclosure norms. This is consistent with the argument that textual complexity represents a low-cost and low-risk strategy to signal ESG commitment, especially when real transformation is financially burdensome.

5.3. Policy implication

This study offers several actionable implications for ESG regulatory practices and corporate reporting.

5.3.1. Integrating readability metrics into ESG disclosure frameworks

Given the strategic use of textual complexity as a signalling mechanism, regulatory bodies should consider incorporating readability standards into ESG reporting guidelines. Tools such as the Fog Index, Flesch-Kincaid can be localized to develop a readability rating system tailored to Chinese ESG contexts [13,14]. Moreover, sector-specific minimum readability thresholds could be introduced to provide soft constraints and issue public alerts when reports fall below acceptable standards.

5.3.2. Implementing sector-specific regulatory responses

The findings suggest that sectors with higher ESG risk profiles are more prone to language-based greenwashing. Regulators should introduce differentiated ESG disclosure benchmarks by sector, especially in industries with greater environmental impacts and higher public visibility.

5.3.3. Enhancing investor capacity to assess ESG text quality

As ESG reports become more complex, traditional rating systems and investor due diligence may be insufficient detect greenwashing. Market actors should be encouraged to adopt textual analytics tools and build literacy around readability metrics. Educational initiatives and open-access training programs can be developed to bridge this knowledge gap.

5.3.4. Developing a national ESG text evaluation platform

Academic institutions, industry associations, and regulatory bodies can collaborate to build a nationwide ESG disclosure assessment system, incorporating dimensions such as readability, consistency, tone, and transparency. This would offer a more holistic and standardized reference for ESG report quality, serving both policymakers and market participants.

5.4. Limitations

While this study provides novel insights into the textual dimensions of greenwashing, it is subject to certain limitations that warrant further exploration. First, the analysis is based on A-share listed companies in China, which may limit generalizability to other institutional contexts with different disclosure regimes. Future studies could adopt a cross-country design to explore whether similar readability patterns exist in markets with varying regulatory intensities.

Second, the RA index used in this study primarily captures syntactic complexity. However, readability is a multidimensional concept influenced by structure, tone, and semantic density. Future work should employ more comprehensive textual analysis methods, including sentiment analysis, topic modeling, and machine learning-based readability metrics.

Last but not least, while the study controls for firm and year fixed effects and conducts robustness checks through subgroup and interaction models, the findings remain correlational. Future research

could incorporate quasi-natural experimental methods such as difference-in-difference (DID) and regression discontinuity design (RDD) to strengthen casual identification and validate the strategic use of language in corporate disclosure behavior.

6. Conclusion

In the context of growing regulatory scrutiny and stakeholder attention to ESG disclosures, this study investigates the relationship between ESG report readability and greenwashing behavior, using a large panel of Chinese A-share listed firms from 2009 to 2023. Based on empirical analysis, the following key findings emerge:

First, there exists a significant positive association between textual complexity and the degree of greenwashing. This suggests that firms may strategically employ obfuscatory language to mask deficiencies in actual ESG performance, thereby engaging in symbolic rather than substantive compliance.

Second, firm size significantly moderates this relationship. In larger firms, the effect of reduced readability on greenwashing is considerably weaker, likely due to stronger external monitoring, greater reputational stakes, and higher compliance expectations. In contrast, smaller firms—often facing resource constraints—are more prone to rely on linguistic complexity as a cost-effective strategy to meet disclosure norms superficially.

Third, substantial industry heterogeneity is observed. The manufacturing sector, which typically faces intense environmental scrutiny yet high compliance costs, exhibits stronger tendencies toward readability manipulation. By contrast, in heavily regulated sectors such as energy, where disclosure standards and public oversight are stricter, the role of textual complexity is less pronounced.

Taken together, the findings advance the literature on greenwashing by integrating a textual perspective and highlight the critical role of language in non-financial corporate reporting. The results also offer important regulatory implications: monitoring ESG disclosure quality should not be limited to content presence but should extend to textual transparency. Specifically, readability could serve as a soft signal for identifying potential greenwashing behavior, especially among small firms in high-risk industries.

Future research should deepen this inquiry by leveraging cross-country comparative datasets, employing machine learning-based text analytics, or examining capital market reactions to textual cues in ESG reporting. In addition, quasi-experimental designs such as difference-in-differences (DID) or regression discontinuity (RDD) could be employed to strengthen causal inference regarding the strategic use of language in sustainability disclosures.

References

[1] M. Spence, Job market signaling. Q. J. Econ. 87, 355 – 374 (1973).

[2] T. P. Lyon, A. W. Montgomery, the means and end of greenwash. Organ. Environ. 28, 223 – 249 (2015). https://www.jstor.org/stable/26164732.

[3] W. Xu, Y. Wu, M. Li, J. Peng, Fishing for the moon in the water or truly effective: ESG ratings and corporate greenwashing risk. Econ. Rev. (02), 132 – 147 (2025). https://doi.org/10.19361/j.er.2025.02.08.

[4] F. Li, Annual report readability, current earnings, and earnings persistence. J. Account. Econ. 45, 221 – 247 (2008).

[5] T. Loughran, B. McDonald, Measuring readability in financial disclosures. J. Finance 69, 1643 – 1671 (2014).

[6] D. Lu, W. Chen, L. Zhang, Readability of financial reports and quality of information disclosure: Empirical evidence from China's capital market. Account. Res. (10), 70 – 78 (2019).

[7] S. Du, C. B. Bhattacharya, S. Sen, Maximizing business returns to corporate social responsibility (CSR): The role of CSR communication. Int. J. Manag. Rev. 12, 8 – 19 (2010). https://doi.org/10.1111/j.1468 - 2370.2009.00276.x.

[8] W. Sun, Board secretary reputation and the readability of information disclosure: Evidence from annual report text mining of A-share companies in Shanghai and Shenzhen. Econ. Manag. 41, 136 – 153 (2019). https://doi.org/10.19616/j.cnki.bmj.2019.07.009.

[9] A. Barnett, K. Leoffler, Readability of accounting and auditing messages. J. Bus. Commun. 16, 49 – 59 (1979).

[10] J. Huang, P. Hu, D. Wang, Y. Wang, The double signal of ESG reports: Readability, growth, and institutional influence on firm value. Sustainability 17, 2514 (2025). https://doi.org/10.3390/su17062514.

[11] B. J. Bushee, I. D. Gow, D. J. Taylor, Linguistic complexity in firm disclosures: Obfuscation or information? J. Account. Res. 56, 85 – 121 (2018).

[12] D. M. Merkl-Davies, N. M. Brennan, Discretionary disclosure strategies in corporate narratives: Incremental information or impression management? J. Account. Lit. 26, 116 – 196 (2007).

[13] R. Gunning, The technique of clear writing (McGraw-Hill, New York, 1952).

[14] J. P. Kincaid, R. P. Fishburne, R. L. Rogers, B.S. Chissom, Derivation of new readability formulas (automated readability index, fog count and Flesch reading ease formula) for Navy enlisted personnel (Research Branch Report 8 – 75, Naval Tech. Train. Command, Millington, 1975).