# Readability of ESG Topic-Specific Sustainability Reports and Bank Loan Spreads

Yun Hao

E-mail: <u>peterhao.mg12@nycu.edu.tw</u> Department of Management Science, National Yang Ming Chiao Tung University, TAIWAN

Hsin-Wei Chen

E-mail: thisiszx0105@gmail.com

Department of Management Science, National Yang Ming Chiao Tung University, TAIWAN

Kai Wang

E-mail: maciek980707@gmail.com

Department of Management Science, National Yang Ming Chiao Tung University, TAIWAN

Yijie Tseng E-mail: <u>081647@mail.fju.edu.tw</u> Department of Accounting, Fu Jen Catholic University, TAIWAN

Tsung-Kang Chen\*

Email: vocterchen@nycu.edu.tw

Department of Management Science, National Yang Ming Chiao Tung University, TAIWAN Center for Research in Econometric Theory and Applications, National Taiwan University, TAIWAN

<sup>\*</sup> Corresponding author. Address: No. 1001, Dasyue Rd., East Dist., Hsinchu City 300, Taiwan (R.O.C.). Tel: +886-3-5712121 ext 57124.

# Readability of ESG Topic-Specific Sustainability Reports and Bank Loan Spreads

## Abstract

We explore the association between the readability of ESG (namely environmental, social, and governance) topic-specific sustainability reports and bank loan spreads using a sample of 4,273 sustainability reports issued by publicly listed companies in Taiwan from 2014 to 2022. Empirical results show that the readability of sustainability reports focused on environmental and social topics are both significantly and negatively related to bank loan spreads when controlling for firm, bank, and loan characteristics variables. We also find that the positive tone of sustainability reports enhances the negative association between the readability of social topic-specific sustainability reports and bank loan spreads while ESG performance weakens the negative association. Besides, these findings hold after addressing endogeneity and sample selection bias issues. Finally, banks are suggested to consider the readability of the E- and S-topic specific sustainability reports in loan practices.

Keywords: ESG topic-specific sustainability report, Readability, bank loan spread, ESG performance

## I. Introduction

In 2004, the United Nations Global Compact collaborated with 20 financial institutions to publish the "Who Cares Wins" report, advocating for integrating environmental, social, and governance (ESG) factors into investment decisions. This report marked the first use of the term "ESG." Since then, climate change has garnered significant attention from governments and the public worldwide, driving a global movement toward sustainable development. In Taiwan, the Financial Supervisory Commission (FSC) mandated in 2023 that listed companies with a capital of over NT\$2 billion must produce sustainability reports, highlighting these as essential tools for stakeholders to assess corporate non-financial information. However, to align with market expectations or secure specific business opportunities and resources, companies may manipulate or obscure the content of sustainability reports to conceal unfavorable information. Such practices could increase risks for lending banks, potentially leading to higher loan spreads. While prior research has predominantly focused on corporate ESG performance, this study focuses on the textual characteristics of ESG topic-specific sustainability reports. Specifically, it examines the association between the readability of sustainability reports across environmental, social, and governance domains and bank loan spreads, aiming to offer recommendations for banks in incorporating sustainability report data into their credit assessments.

Although numerous studies have explored ESG performance metrics, few have analyzed the textual features of sustainability reports. For instance, Yu and Garg (2022) observed that companies with less readable corporate social responsibility (CSR) reports tend to have higher default risks. Similarly, other studies have centered on environmental performance indicators (Lee et al., 2016) or ESG ratings (Duque-Grisales & Aguilera-Caracuel, 2021), with relatively little attention given to textual features like ESG-related keywords, sentiment, tone, or readability. Furthermore, there is a lack of research disaggregating textual analyses into the distinct E, S, and G categories.

To classify the themes within sustainability reports, this study utilizes the ESG wordlist

developed by Baier et al. (2020) and the ANTUSD sentiment dictionary proposed by Wang and Ku (2016). The primary ESG theme of each report is determined by analyzing the proportion of environmental (E), social (S), and governance (G) keywords. Subsequently, the sentiment tone of the text is evaluated to gauge its positivity, enabling a more profound analysis across various subcategories. Regarding textual features, this study focuses on readability as the independent variable, while keyword proportions and sentiment tone are used in the categorization process. By combining text classification with diverse textual characteristics, the study aims to provide external stakeholders with a clearer understanding of companies' emphasis on the E, S, and G domains and the communicative value of their sustainability reports.

Bank loan spreads were selected as the research variable for two main reasons. First, among various stakeholders, banks have been at the forefront of ESG-related initiatives, such as the Equator Principles and Taiwan's Green Finance Action Plan 3.0, making them more proactive in implementing ESG practices. Second, bank loans represent the most common form of corporate financing, incentivizing companies to craft external disclosures to secure lower loan rates carefully. This study, therefore, seeks to analyze whether the relationship between sustainability report readability and bank loan spreads varies based on a company's focus on different ESG themes. Doing so aims to help financial institutions identify potential manipulations in sustainability report texts.

Building on the potential for report manipulation and its implications for lending decisions, this study draws on Bloomfield (2002) management obfuscation hypothesis and Duffie and Lando (2001) incomplete accounting information credit risk model. The study examines the relationship between sustainability report readability and bank loan spreads from an information asymmetry perspective. According to these hypotheses, poorly performing companies often reduce the readability of their reports to obscure unfavorable information. As such, this study hypothesizes a negative relationship between sustainability report readability and bank loan spreads. Furthermore, the study incorporates positive sentiment tone and ESG performance as moderating variables. It predicts that higher levels of positive sentiment will amplify the negative relationship between readability and loan spreads, while stronger ESG performance will mitigate this relationship.

Empirical results indicate that the readability of sustainability reports focusing on environmental (E-Topic) and social (S-Topic) aspects exhibits a significant negative relationship with bank loan spreads, supporting the main hypothesis. The moderation model further reveals that high positive sentiment within the text enhances the importance of readability, significantly strengthening the negative relationship between the readability of social-themed sustainability reports and bank loan spreads. Conversely, superior ESG performance reduces banks' concerns regarding information transparency, weakening this negative relationship. To address potential concerns about sample selection bias, the study employs the Heckman two-stage sample selection model, with results confirming that the main findings remain robust after correcting for bias. Moreover, instrumental variable regression is utilized to address endogeneity issues, and the results affirm the validity of the conclusions.

Compared to prior studies, such as Yu and Garg (2022), which examined the relationship between Corporate Social Responsibility (CSR) report readability and bank loan spreads, this study extends the scope by categorizing sustainability reports into four distinct classifications: environment-focused, social-focused, governance-focused, and integrated ESG perspectives. This classification facilitates a deeper analysis of how the readability of reports focusing on different topics correlates with bank loan spreads, providing nuanced insights into the communication value of each type of sustainability report. Furthermore, this study incorporates sentiment tone analysis, distinguishing reports based on high and low levels of positive sentiment. It investigates whether company managers strategically manipulate report readability to emphasize positive information or obscure unfavorable news, consistent with obfuscation strategies. The findings suggest practical implications for financial institutions, recommending that banks consider the readability, thematic focus, sentiment tone, and ESG performance of sustainability reports during credit assessment processes. Such considerations can aid in mitigating lending risks while aligning with green finance objectives and maximizing the benefits of ESG strategy integration in banking practices.

## **II. Literature Review and Hypothesis Development**

This section is divided into four parts: (1) Types of textual features and related literature, (2) studies related to bank loan spreads, (3) research and theories connecting sustainability report textual features with bank loan spreads, and (4) moderating effects and their related literature, followed by hypotheses development.

## II.1. Textual Features

Corporate reports, typically consisting of quantitative and qualitative content, offer stakeholders intuitively accessible information and are frequently utilized as research variables. While quantitative data is more structured and straightforward to analyze, qualitative content, such as shareholder reports or sustainability strategy blueprints, presents more significant analytical challenges due to its unstructured nature. Qualitative data often contain more prosperous, more meaningful insights (De Pelsmacker et al., 2018). Consequently, an increasing number of scholars have adopted computational methods in recent years to transform textual content into quantifiable data, thereby generating various textual features that enable more in-depth research on unstructured texts. This study leverages these methodologies to analyze the textual features of sustainability reports.

In prior research focusing on annual reports, financial statements, and shareholder reports, commonly studied textual features include readability, tone (or sentiment), and domain-specific keywords. However, these textual features have not been comprehensively defined or extensively applied in the context of Corporate Social Responsibility (CSR) or Environmental, Social, and Governance (ESG) reports. This study addresses this gap by integrating three key textual features. First, ESG-related keywords and sentiment tones cluster sustainability report samples. Subsequently, readability is used as the primary independent variable for analysis. The subsequent sections detail the relevant literature on readability, sentiment tone, and domain-specific keywords:

## II.1.1. Readability

Readability refers to the ease with which a text can be read and understood (Klare, 1963). Texts that are simpler to comprehend are characterized by higher readability. Key factors influencing readability include writing style, text organization, and content (Dale & Chall, 1949). Contemporary readability indices are typically calculated using average sentence length, the number of complex words, total word count, and sentence count.

In financial text readability research, various studies have explored its implications for corporate reporting and stakeholder communication. For instance, Lo et al. (2017) employed the Fog Index to assess the readability of annual reports' Management Discussion and Analysis (MD&A) sections, concluding that companies producing more complex MD&A content are more likely to engage in earnings management. Sun et al. (2022) examined the impact of adopting the International Financial Reporting Standards (IFRS) on the readability of Chinese corporate reports, finding that while these reports became more extended, their readability improved. Similarly, Jiang (2023) measured readability based on the prevalence of difficult words in CSR reports and discovered that Taiwanese companies with superior CSR performance tended to use more complex vocabulary. This suggests that the inclusion of such words reflects substantive meaning rather than mere complexity or obscurity.

#### II.1.2. Sentiment

Sentiment tone refers to the emotional or opinion-oriented content expressed within a text. The most common sentiment indicators used in research measure the positive or negative tone of the text(Agarwal et al., 2016). This study focuses on positive sentiment as the primary measure of sentiment tone. Other sentiment indicators, such as extremity (Bochkay et al., 2020) and euphemism(Suslava, 2021), have also been explored in prior literature.

The dictionary method is widely employed in financial text analysis to evaluate sentiment tone. For instance, Rogers et al. (2011) utilized the Loughran-McDonald Master Dictionary (LM Dictionary) to analyze earnings announcements, finding that texts with a more optimistic tone were associated with higher litigation risks. Similarly, Price et al. (2012) applied the Harvard Psychological Dictionary and the Henry Dictionary (Henry, 2006) to measure sentiment tone in earnings call transcripts, discovering that the tone could predict abnormal returns and trading volumes. Furthermore, they highlighted the additional value of the Q&A sections of earnings calls for external stakeholders, as these sections often provided richer insights. This study adopts a similar dictionary-based approach to measure sentiment tone in sustainability reports. Specifically, it employs the augmented Chinese opinion lexicon (ANTUSD) to calculate sentiment scores, enabling a systematic evaluation of positive sentiment within the texts. This approach aligns with the broader objective of exploring how sentiment tone contributes to understanding the communication value of sustainability reports.

## II.1.3. Domain-Specific Keywords

Keywords, such as sentiment tone, represent textual features that encapsulate the meaning conveyed by a given text. The keywords analysis facilitates understanding of a company's emphasis on various issues and identifies potential practices such as the concealment of harmful information or management obfuscation. Nevertheless, keyword analysis relies heavily on preexisting domain-specific dictionaries, which has limited its widespread adoption compared to other textual feature analyses.

In the context of financial text analysis, Loughran (2019) compiled a list of 130 oil-related keywords. They demonstrated that terms like "output cut" and "demand up" in news articles predicted a decline in oil prices on the subsequent trading day, thereby illustrating the price correction process following market overreaction. Similarly, Heichl and Hirsch (2023), utilizing the ESG dictionary developed by Baier et al. (2020), discovered that Swedish companies, driven by progressive net-zero carbon policies, exhibited a higher prevalence of environmental keywords in their sustainability reports compared to companies in the United Kingdom, France, and Italy.

This study builds on the work of Heichl and Hirsch (2023) by employing the ESG dictionary proposed by Baier et al. (2020) to identify ESG-related topics. To the best of the authors' knowledge, this is the first investigation to analyze the ESG keywords within Chinese

sustainability reports while also exploring the relationship between the readability of topicspecific sustainability reports and bank loan spreads.

#### II.2. Studies on Bank Loan Spreads

Hrazdil et al. (2024) highlight that companies with more frequent negative climate reputation events tend to experience higher loan spreads, face stricter contractual restrictions, and have shorter borrowing terms. These findings suggest that banks increasingly incorporate climate risk into their lending and investment decisions. Similarly, Yu and Garg (2022), adopting the research methodology of Loughran and McDonald (2014), use file size as a proxy for text readability and demonstrate that companies with less readable CSR reports are associated with higher default risks and elevated bank loan spreads. They argue that report readability serves as an effective channel for conveying company value information.Building on this inference, this study investigates the relationship between the readability of sustainability reports across different topics and bank loan spreads. However, the methodology employed in this paper differs from that of Yu and Garg (2022). Specifically, this study categorizes the sustainability reports of publicly listed companies in Taiwan into four primary categories based on the ESG keyword dictionary developed by Baier et al. (2020): environmental focus (E-Topic), social focus (S-Topic), corporate governance focus (G-Topic), and comprehensive ESG themes (ESG-Topic). By analyzing the variations in the relationships between the readability of sustainability reports across these ESG topics and bank loan spreads, this paper aims to provide deeper insights into the communicative value of different types of sustainability reports.

# II.3. Research on the Relationship Between Sustainability Report Readability and Bank Loan Spreads

From the preceding discussion, it can be inferred that low readability in a company's publicly available reports may signal an intent to obscure information. This interpretation aligns with Courtis (1998) and the Obfuscation Hypothesis, which posits that companies with lower readability or more significant variability in the readability of their annual reports are more

likely to associate unfavorable news with high-level coverage, thereby reducing the clarity of information and confusing the public. Expanding upon this, Bloomfield (2002) proposed the Incomplete Revelation Hypothesis, which asserts that when a company performs poorly, managers may intentionally complicate or obscure information to conceal adverse outcomes an idea referred to as the Management Obfuscation Hypothesis. Within this theoretical framework, this study posits that reduced readability reflects a greater prevalence of complex language and higher textual intricacy, thereby exacerbating information asymmetry between the company and its external stakeholders.

In examining the relationship between information asymmetry and bank loans, this study draws upon the Incomplete Accounting Information credit risk model proposed by Duffie and Lando (2001). This model suggests that incomplete disclosure of corporate information heightens credit risk, with greater information asymmetry leading to higher credit spreads. Empirical evidence supports this perspective; for instance, Yu (2005) demonstrated that firms with higher transparency in accounting information tend to have lower credit spreads. Raimo observed that companies with more comprehensive ESG disclosures achieve improved financing conditions. Consequently, low readability or ambiguous language often indicates Managerial Information Hoarding, resulting in diminished transparency and increased external financing costs (Ertugrul et al., 2017).

Conversely, the social dimension of corporate social responsibility (CSR) extends its influence across a broader spectrum of stakeholders, enabling social ESG information to significantly impact a company's economic performance (Sila & Cek, 2017). This information also directly shapes the perceptions of customers, the public, and suppliers (Iamandi et al., 2019). As such, higher readability in sustainability reports focusing on social aspects facilitates the effective communication of relevant social issues, reducing lending risks for banks and enhancing their willingness to offer more favorable loan terms and lower interest rates.

In conclusion, the readability of sustainability reports is expected to exhibit a negative relationship with bank loan spreads, with reports emphasizing social aspects having a more pronounced effect. Based on these insights, the study proposes the following hypotheses1& hypotheses 2:

H1: The readability of sustainability reports, focused on each ESG dimension (environmental, social, governance, and comprehensive ESG), has a negative relationship with bank loan spreads.

**H2**: The negative relationship between the readability of sustainability reports and bank loan spreads is stronger for reports focused on social aspects than the other three dimensions (environmental, governance, and comprehensive ESG).

II.4. The Moderating Effect of Positive Textual Tone and ESG Performance

# II.4.1. The Moderating Effect of High Positive Tone on the Relationship Between Sustainability Report Readability and Bank Loan Spreads

As discussed previously, the readability of sustainability reports is anticipated to exhibit a negative relationship with bank loan spreads. Beyond readability, however, the tone of a text constitutes another critical textual feature. Optimistic language in a company's reports often reflects positive expectations for the future. For instance, Davis et al. (2012) demonstrated that the positivity of a company's revenue press releases signals management's confidence in the company's future performance, revealing a positive relationship between the tone of the press release and the company's subsequent return on assets (ROA). Similarly, Tsai & Wang(2017) identified negative language in financial reports as a significant predictor of future economic risks.

Building on these findings, this study proposes that the tone of sustainability reports should exert a comparable influence. Specifically, a positive tone is expected to moderate the relationship between readability and bank loan spreads. A highly positive tone in sustainability reports may strengthen the negative relationship between readability and bank loan spreads by enhancing stakeholders' perception of the company's commitment to ESG practices. Based on this reasoning, the study formulates the following hypothesis:

H3: A highly positive tone amplifies the negative relationship between the readability of

sustainability reports, focused on each ESG dimension, and bank loan spreads.

II.4.2. The Moderating Effect of Good ESG Performance on the Relationship Between Sustainability Report Readability and Bank Loan Spreads

ESG or CSR performance reflects a company's non-financial achievements and is a crucial tool for external stakeholders to evaluate non-financial information. Enhanced ESG performance signifies a company's accomplishments beyond financial metrics, reducing information asymmetry between the company and its stakeholders (Cui et al., 2018). Based on this premise, the study argues that better ESG performance by a borrowing company may lead lending banks to reduce their reliance on the content of sustainability reports during the credit assessment process. Consequently, the significance of sustainability report readability in influencing lending decisions diminishes.

Building on this reasoning, the study hypothesizes that higher ESG performance weakens the importance of sustainability report readability in determining bank loan spreads. Specifically, improved ESG performance may reduce banks' concerns about information transparency, thereby mitigating the impact of report readability on credit risk assessments. **H4:** Good ESG performance will weaken the negative relationship between the readability of sustainability reports focused on each ESG dimension and bank loan spreads.

## **III. Data and Methodology**

## III.1. Data Sources and Sample Selection

This study uses sustainability reports from publicly listed companies in Taiwan as the sample, covering the period from 2014 to 2022. The primary data source is the Market Observation Post System (MOPS), which includes 4,693 sustainability reports publicly declared during this period. After excluding reports with missing files for specific years and those that could not be successfully converted into text files, the final sample consists of 4,273 reports. The distribution of samples across the years is detailed in Column (5) of Table 1. For

ESG topic classification, this study utilizes the ESG Wordlist proposed by Baier et al. (2020) to conduct keyword analysis and text classification, with the results summarized in Table 1.

## [Insert Table 1 here]

To calculate the emotional tone of the sustainability reports, the study employs the Augmented NTU Sentiment Dictionary (ANTUSD) developed by Wang and Ku (2016) for sentiment analysis. The readability of the reports, serving as the independent variable, is measured using the Chinese Readability Index Explorer (CRIE) developed by Sung et al. (2016). Dependent variables related to corporate loans, moderating variables, financial characteristics, and control variables for loan attributes are obtained from the Taiwan Economic Journal (TEJ) database.

This comprehensive data collection and methodology enable a robust analysis of the relationship between sustainability report readability and bank loan spreads while accounting for ESG topic classification, sentiment tone, and relevant financial characteristics.

#### III.2. Variable Definitions

#### III.2.1. Measurement and Classification of Sustainability Report Text Topics

This study adopts the ESG topic classification approach outlined by Heichl and Hirsch (2023), employing the ESG dictionary Baier et al. (2020) developed to calculate the frequency and proportion of ESG-related keywords for topic identification. A sustainability report is classified into a specific ESG topic based on the proportion of keywords in the text, surpassing the median for the entire sample. For instance, if the proportion of environmental (E) keywords in a report exceeds the median, the report is classified as having an environmental focus (E-Topic). The same criterion is applied to social (S-Topic) and governance (G-Topic) classifications. Appendix Table 1 provides a subset of the translated ESG dictionary used in this analysis.

Additionally, reports classified simultaneously as E-Topic, S-Topic, and G-Topic are categorized as encompassing all three ESG dimensions (ESG-Topic). It is important to note that

these classifications are not mutually exclusive. A single sustainability report can be associated with multiple categories or none, depending on its content and alignment with the criteria for each ESG dimension. This flexible classification approach allows for a nuanced analysis of sustainability report themes and their relationship with bank loan spreads.

## III.2.2. Measurement of Text Readability

This study draws on the Dale-Chall Formula introduced by Dale (1995), the Fog Index employed by Li (2008), and the Chinese readability formulas developed by Sung et al. (2016). Chinese and English readability formulas consider the number or proportion of complex words as a critical variable influencing readability. Recognizing that the total word count of sustainability reports can vary significantly across industries, this study adopts the proportion of complex words as a substitute variable, consistent with the approach used by Ertugrul et al. (2017). This proportion, referred to as *COMPLEX\_WORDS\_PERCENT*, is calculated by dividing the number of complex words by the total word count of the text.

To measure the number of complex words, this study utilizes the Chinese Readability Index Explorer (CRIE) system developed by Sung et al. (2016). Complex words are those not included in the top 3,000 most frequently used words in the Academia Sinica Balanced Corpus. By employing this refined metric, the study ensures an accurate and consistent assessment of text readability across sustainability reports, enabling a robust analysis of its relationship with bank loan spreads.

# $COMPLEX_WORDS_PERCENT = \frac{Complex Words}{Total Words}$

## III.2.3. Measurement of Text Sentiment

This study employs the Augmented Chinese Sentiment Dictionary (ANTUSD) developed by Wang and Ku (2016) to calculate sentiment scores for the text. Each word in the dictionary is assigned a sentiment score ranging from -1 to +1, with scores less than zero indicating negative sentiment, zero representing neutral sentiment, and scores greater than zero indicating positive sentiment. Following the methodology outlined by Kuo et al. (2021), the text is first segmented using the Jieba Chinese segmentation tool. Subsequently, the overall sentiment score is calculated using the *ANTUSD* dictionary and normalized by dividing it by the total word count of the text, yielding the *TONE\_SCORE*, as expressed in the formula below:

$$TONE\_SCORE = \frac{Total \ Tone \ Score}{Total \ Words}$$

After computing the sentiment scores for all report samples, the results indicate that all reports exhibit a positive sentiment. This outcome aligns with prior research, suggesting that sustainability reports generally emphasize positive information and that larger companies avoid negative wording (Mućko, 2021). To facilitate further analysis, texts with sentiment scores above the median are classified as having a high positive tone (High). In contrast, those below the median are classified as having a low positive tone (Low). A detailed summary of this classification is presented in Table 2.

## [Insert Table 2 here]

#### III.2.4. Measurement of Bank Loan Spreads

This study adopts the methodology proposed by Hrazdil et al. (2024) to calculate bank loan spreads (*LOAN\_SPREAD*). The calculation is based on the difference between the loan interest rate and the Taipei Interbank Offered Rate (*TAIBOR*) which helps eliminate the impact of benchmark rate fluctuations caused by varying economic conditions.

The loan interest rate is determined as the average of the highest and lowest interest rates for each loan, with data sourced from the Taiwan Economic Journal (TEJ) database, specifically the banking loan details for publicly listed companies in Taiwan. The *TAIBOR* used in this calculation corresponds to the monthly average for the period in which the loan data is recorded and is obtained from the Bankers Association of the Republic of China.

This approach ensures that *LOAN\_SPREAD* reflects the company-specific risk premiums and credit terms, enabling a robust analysis of its relationship with sustainability report characteristics.

### III.3. Control Variables

Building upon prior studies that demonstrate a negative correlation between ESG performance and default risk (Atif & Ali, 2021) and between CSR implementation and credit risk (Hunjra et al., 2024), this study adopts a regression model referencing Yu and Garg (2022). To ensure robustness, control variables are divided into company and loan characteristics while accounting for fixed effects related to the year, company, and lending bank.

In terms of company characteristics, this study controls for firm size (*FIRM\_SIZE*), financial leverage(*LEVERAGE*), market-to-book ratio (*MB*), profitability (*PROFITABILITY*), asset tangibility (*TANGIBILITY*), Altman's Z-score bankruptcy indicator (*BANKRUPT*), credit risk indicator (*TCRI*), ESG performance (*ESGP*), and report sentiment (*TONE\_SCORE*). *FIRM\_SIZE* is defined as the natural logarithm of total assets; *LEVERAGE* is the ratio of long-term debt to total assets; *MB* is the ratio of ending common stock market value to total equity; *PROFITABILITY* is EBITDA divided by total assets; *TANGIBILITY* is the book value of long-term real estate, plants, and equipment divided by total assets; *BANKRUPT* is measured using Altman (2000) to measure a company's bankruptcy risk; *TCRI* uses the total score of sustainability development indicators (TESG) from TEJ; and *TONE\_SCORE* is the average sentiment score calculated by this study.

For loan characteristics, the study controls for loan size (*LOAN\_SIZE*), loan term (*MATURITY*), and secured loans (*SECURITY*). *LOAN\_SIZE* is defined as the natural logarithm of the total loan amount; *MATURITY* is the loan term in years; and *SECURITY* is a dummy variable, with secured loans coded as one and unsecured loans as 0.

## III.4. Model Development

The methodology for hypothesis testing involves three regression models, each tailored to validate specific hypotheses. Given that the dataset comprises panel data from multiple companies over several years, the models incorporate a lagged effect of sustainability report readability to account for timing differences in report submissions. Fixed effects for a year, company, and lending bank are included to control for unobservable factors, and residual heteroscedasticity is addressed at the loan level to ensure robust results.

#### III.4.1. Main Model

The main model is designed to test Hypotheses H1 and H2. It first investigates whether the readability of sustainability reports focused on different ESG dimensions is negatively related to bank loan spreads, as posited by H1. Subsequently, it conducts a coefficient difference test to evaluate whether the negative relationship between the readability of socially-focused sustainability reports (S-Topic) and bank loan spreads is stronger compared to reports centered on environmental (E-Topic), governance (G-Topic), and comprehensive ESG (ESG-Topic) dimensions. This step validates Hypothesis H2.

$$\begin{aligned} LOAN\_SPREAD_{i,j,t+1} &= \beta_0 + \beta_1 \ COMPLEX\_WORDS\_PERCENT_{i,t} + \beta_2 \ FIRM\_SIZE_{i,t} \\ &+ \beta_3 \ LEVERAGE_{i,t} + \beta_4 \ MB_{i,t} + \beta_5 \ PROFITABILITY_{i,t} \\ &+ \beta_6 \ TANGIBILITY_{i,t} + \beta_7 \ BANKRUPT_{i,t} + \beta_8 \ TCRI_{i,t} + \beta_9 \ ESGP_{i,t} \\ &+ \beta_{10} \ TONE\_SCORE_{i,t} + \beta_{11} \ LOAN\_SIZE_{i,j,t} + \beta_{12} \ MATURITY_{i,j,t} \\ &+ \beta_{13} \ SECURITY_{i,j,t} + YEAR\_FE + FIRM\_FE + BANK\_FE \\ &+ \varepsilon_{i,j,t+1} \end{aligned}$$
(1)

Here, *i*, *j*, and *t* represent the loan *j* for company *i* in year *t*. A higher proportion of complex words (*COMPLEX\_WORDS\_PERCENT*) indicates lower readability, leading to higher bank loan interest rates. Therefore, this study expects that  $\beta_1$  in Eq. (1) will be positively significant across all subsamples and that the coefficient difference test for the model focused on the social dimension will be significant compared to the other subsamples

## III.4.2. Moderating Effect Model

The moderating effect model tests hypotheses H3 and H4, investigating whether the relationship between the readability of sustainability reports and bank loan spreads is moderated by high positive tone texts and ESG performance. To validate hypothesis H3— the moderating effect of text positivity—high positive tone text is treated as a categorical variable. Thus, this

study employs a sample grouping method to test the moderating effect on the regression equation (1). For hypothesis H4, which examines the moderating effect of ESG performance, ESG performance is treated as a continuous variable, and this study employs a regression model that incorporates interaction terms (as in regression equation (2)).

$$\begin{aligned} LOAN\_SPREAD_{i,j,t+1} &= \beta_0 + \beta_1 \ COMPLEX\_WORDS\_PERCENT_{i,t} + \beta_2 \ ESGP_{i,t} \\ &+ \beta_3 \ (COMPLEX\_WORDS\_PERCENT_{i,t} \times ESGP_{i,t}) + \beta_4 \ FIRM\_SIZE_{i,t} \\ &+ \beta_5 \ LEVERAGE_{i,t} + \beta_6 \ MB_{i,t} + \beta_7 \ PROFITABILITY_{i,t} \\ &+ \beta_8 \ TANGIBILITY_{i,t} + \beta_9 \ BANKRUPT_{i,t} + \beta_{10} \ TCRI_{i,t} \\ &+ \beta_{11} \ TONE\_SCORE_{i,t} + \beta_{12} \ LOAN\_SIZE_{i,j,t+1} + \beta_{13} \ MATURITY_{i,j,t+1} \\ &+ \beta_{14} \ SECURITY_{i,j,t+1} + YEAR\_FE + FIRM\_FE + BANK\_FE \\ &+ \varepsilon_{i,j,t+1} \end{aligned}$$
(2)

Among them, i, j and t represent the loan j of company i in year t. The expected results are described below:

In hypothesis H3, the positive tone of sustainability reports is expected to reflect a company's optimistic outlook for the future, enhancing the effectiveness of information transmission related to text readability. Therefore, this study posits that highly positive texts will strengthen the moderating effect. It is anticipated that in the subsample using high positive tone texts, the coefficient  $\beta_1$  in regression equation (1), will be significantly positive and more significant than the regression coefficient for the low positive tone subsample.

For hypothesis H4, since better ESG performance (ESGP) indicates strong non-financial performance and reduces information asymmetry, the importance of sustainability report readability is expected to diminish. Consequently, this study believes ESG performance will weaken the negative relationship between sustainability report readability and bank loan spreads. It is anticipated that in regression equation (2),  $\beta_1$  will be positive, while  $\beta_3$  will be significantly negative.

## **IV. Research Results**

IV.1. Descriptive Statistics and Correlation Analysis

This study includes 13,769 observations for bank loan spreads, where each company may have multiple loans in a given year, resulting in several observations exceeding that of the sustainability report samples. Descriptive statistics for the study variables are presented in Table 3. For the dependent variable, the average bank loan spread (*LOAN\_SPREAD*) is 0.472%, with a standard deviation of 0.594%, indicating considerable variation in loan costs across companies, which necessitates the use of winsorization. Regarding textual features, the average percentage of complex words (*COMPLEX\_WORD\_PERCNT*) is 33.9%, with a standard deviation of 2.3%. The average sentiment score (*TONE\_SCORE*) is 0.123, with a standard deviation of 0.013, suggesting that most of the samples exhibit a positive tone.

## [Insert Table 3 here]

In addition, Pearson correlation coefficients are employed to assess the relationships between the variables. The correlation coefficients for all primary variables are calculated and displayed in Table 4. Notably, the bankruptcy indicator(BANKRUPT) and the market-to-book ratio (MB) show a correlation coefficient greater than 0.7, whereas the majority of other variables exhibit coefficients below 0.3. This indicates that the variables are only weakly correlated, suggesting that multicollinearity is not a concern within the research model.

## [Insert Table 4 here]

#### IV.2. Regression Analysis

This section validates hypotheses H1 and H2 by examining whether the readability of sustainability reports across various dimensions is negatively related to bank loan spreads. It also explores whether the social dimension shows a stronger negative relationship. The empirical analysis results for the primary model are presented in Table 5. After controlling for fixed effects of year, company, and the lending bank, the regression coefficients for the percentage of complex words (*COMPLEX\_WORDS\_PERCENT*) in columns (1), (2), and (3) are 1.409, 1.965, and 2.158, respectively, all significant at the 1% level. These results indicate

a significant positive relationship between the percentage of complex words and bank loan spreads, suggesting that higher complex word percentages correspond to lower readability. The regression findings confirm a significant negative relationship between sustainability report readability and bank loan spreads for the environmental (E-Topic) and social dimension (S-Topic) subsamples, supporting hypothesis H1.

Additionally, the coefficient difference test reveals that the negative relationship between the readability of sustainability reports focused on the social dimension (S-Topic) and bank loan spreads are more robust compared to the governance (G-Topic) and comprehensive ESG (ESG-Topic) subsamples. This is evidenced by the Chi-square statistics from the coefficient difference test in Table 5, which yield values of 11.33 and 10.88, both significant at the 1% level. However, the coefficient difference test between the social (S-Topic) and environmental (E-Topic) texts is insignificant, providing partial support for hypothesis H2.

In summary, through regression equation (1), this study finds that companies with higher percentages of complex words in their sustainability reports focused on environmental (E-Topic) and social dimensions (S-Topic) are likely engaging in managerial information hoarding, which leads to increased information asymmetry with external stakeholders. This, in turn, results in higher loan costs, partially supporting hypothesis H1. Furthermore, sustainability reports focused on the social dimension (S-Topic), which is likely to impact stakeholder perceptions and future revenue performance significantly, strengthening the negative relationship between readability and bank loan spreads compared to governance and comprehensive ESG themes, thus supporting hypothesis H2.

## [Insert Table 5 here]

## IV.3. Robustness Analysis

Although this study includes a range of control variables, the robustness of the model may still be influenced by potential endogeneity concerns. To address these issues, this study employs the average percentage of complex words (*MEAN\_COMPLEX\_WORDS\_PERCENT*) from sustainability reports of companies within the same corporate group for the same year as an instrumental variable. Data on corporate group composition is sourced from the Taiwan Economic Journal (TEJ). Companies within the same group will likely share similar sustainability strategies or face comparable corporate social responsibility pressures. As a result, the management preparing the sustainability reports or the certifying entities may be the same. Consequently, the vocabulary and formatting used in the sustainability reports of companies within the same group are expected to be similar, thereby satisfying the relevance condition for instrumental variables. Moreover, the impact of the average percentage of complex words in group sustainability reports on the financing costs of an individual company is relatively small, fulfilling the exclusion condition for instrumental variables.

The regression results for the instrumental variables are presented in Table 6. Columns (1) and (3) show the results for the subsample focused on the environmental dimension (E-Topic), while columns (2) and (4) present the results for the subsample concentrated on the social dimension (S-Topic). In the first stage, the instrumental variable regression coefficients are all significantly positive, with the F-statistics in columns (1) and (2) exceeding 10, suggesting that the average percentage of complex words in group companies is a valid instrumental variable. In the second stage, the regression coefficients for *MEAN\_COMPLEX\_WORDS\_PERCENT* are positive and significant at the 1% and 5% levels, respectively. These results demonstrate that the rate of complex words continues to affect bank loans significantly spreads for both E-Topic and S-Topic sustainability reports, even after addressing endogeneity concerns.

## [Insert Table 6 here]

Moreover, the sample used in this study consists of companies that have publicly disclosed their sustainability reports. However, companies that must disclose or voluntarily publish sustainability reports are likely to be larger, have better financial performance, or exhibit superior ESG performance, which could introduce sample selection bias. This study employs the Heckman two-step sample selection model for correction to mitigate this concern. The results are shown in Table 7, where the regression coefficients in columns (2), (3), and (4) are all significantly positive. Notably, compared to the regression coefficients from the primary model in Table 5, the coefficient for the social dimension (S-Topic) sustainability report sample increases to 2.205 after correcting for sample selection bias, surpassing the original coefficient of 2.158. Furthermore, the t-value is higher, indicating a stronger negative relationship between the independent and dependent variables.

## [Insert Table 7 here]

In conclusion, regardless of the overall sample or the sustainability report samples focused on the environmental (E-Topic) or social (S-Topic) dimensions, using instrumental variable two-stage regression and sample selection correction models confirms the robustness of the primary regression model's results. This holds even after addressing potential endogeneity and sample selection bias issues.

## IV.4. Moderating Effects Analysis

## *IV.4.1. The Moderating Effect of Positive Tone Text on the Relationship between Sustainability Report Readability and Bank Loan Spreads*

Given that the primary model regression results indicate a negative correlation between sustainability report readability and bank loan spreads for both the E-Topic and S-Topic samples, this study analyzes the moderating effect of positive tone in validating hypothesis H3. The analysis focuses specifically on the E-Topic and S-Topic samples, using a subsampling approach to examine the moderating effect of positive tone. Texts with tone scores above the median are categorized as having a high positive tone (High), while those below the median have a low positive tone (Low). The regression results are presented in Table 8, controlling for fixed effects of year, company, and lending bank.

Columns (3) and (4) of Table 8 show the results for the S-Topic subsample. In column (3),

the regression results for the high positive tone subsample reveal a coefficient for *COMPLEX\_WORDS\_PERCENT* of 6.549, which is significantly positive. In contrast, column (4) shows that for the low positive tone subsample, the coefficient is positive but insignificant. These findings suggest that when the sustainability report exhibits a highly positive tone, the negative relationship between readability and bank loan spreads becomes stronger. However, no moderating effect of high positive tone is observed in the regression results for the E-Topic subsample presented in columns (1) and (2).

## [Insert Table 8 here]

The regression results indicate that when a highly positive tone characterizes the text, it enhances the negative relationship between the readability of sustainability reports and bank loan spreads. This suggests that when the language is more optimistic, it conveys more effective positive information in high-readability texts, thus reducing bank loan spreads. On the other hand, when a highly positive tone characterizes the text, stakeholders may encounter a mixture of complex and positive words in low-readability texts. This combination could raise concerns about managerial obfuscation or the concealment of negative information, increasing information asymmetry. As a result, higher loan spreads are observed, thereby supporting hypothesis H3.

# *IV.4.2. The Moderating Effect of ESG Performance on the Relationship between Sustainability Report Readability and Bank Loan Spreads*

Regarding the moderating effect of ESG performance, this study focuses on the significant E-Topic and S-Topic subsamples identified in the primary model. Since ESG performance is treated as a continuous variable, hypothesis H4 is tested by incorporating interaction terms into the regression model to examine its moderating effect. The regression results, including these interaction terms, are presented in Table 9.

In column (2), which focuses on the social dimension (S-Topic), the coefficient for the

independent variable *COMPLEX\_WORDS\_PERCENT* is 12.382, significantly positive, while the coefficient for the interaction term (*COMPLEX\_WORDS\_PERCENT* × *ESGP*) is -0.161, significantly negative. This indicates that ESG performance weakens the negative relationship between sustainability report readability and bank loan spreads. However, in column (1), focusing on the environmental dimension (E-Topic), although the coefficient for the percentage of complex words is negative, the interaction term's coefficient is significantly positive, demonstrating that the impact of ESG performance on the relationship between complex word usage and bank loan spreads differs within the environmental context.

In summary, when a company's ESG performance is strong, it can effectively communicate sufficient information regarding environmental, social, and governance aspects to lending banks, thereby reducing information asymmetry. As a result, lenders may feel less inclined to scrutinize the sustainability reports closely, diminishing the importance of readability in these reports. Therefore, ESG performance weakens the negative relationship between sustainability report readability and bank loan spreads. However, this moderating effect is only observed in sustainability reports focusing on the social dimension (S-Topic), which supports hypothesis H4.

[Insert Table 9 here]

## **V.** Conclusion

This study explores the relationship between the readability of ESG topic-specific sustainability report and bank loan spreads for publicly listed companies in Taiwan, covering the period from 2014 to 2022. Readability is measured using the proportion of difficult words. Unlike previous literature, this study employs the ESG keyword dictionary proposed by Baier et al. (2020). It categorizes sustainability reports into four types: those primarily focused on environmental aspects (E-Topic), social aspects (S-Topic), governance aspects (G-Topic), and an integrated ESG topic (ESG-Topic). This study further investigates whether companies

emphasizing different ESG themes in their sustainability reports exhibit different association between the readability of sustainability report and bank loan spreads. The main findings confirm a negative relationship between the readability of the E-Topic and S-Topic specific sustainability reports and corresponding bank loan spreads. This indicates that when information is easier to understand, it reduces information asymmetry between the company and external stakeholders, leading to lower bank loan rates. Furthermore, the results show that the negative relationship between the readability of sustainability reports with a social focus (S-Topic) and bank loan spreads is more pronounced than that of governance-focused reports (G-Topic). This suggests that information related to social themes more directly impacts the perceptions of clients, the public, and suppliers regarding the company, thereby influencing operational and financial conditions and leading banks to pay more attention to social content when reviewing sustainability reports.

Regarding the moderating effects, the empirical results reveal that tone positivity enhances the negative relationship between the readability of sustainability reports focused on social aspects (S-Topic) and bank loan spreads. This implies that a positive tone reflects management's confidence in the company's future performance, resulting in lower bank loan spreads for highpositive-tone and high-readability texts. Conversely, in cases where readability is low but the tone is highly positive, it may indicate the use of complex positive language to obscure negative information, reflecting a motivation for managerial obfuscation. This could lead to even higher loan spreads under conditions of low readability with a high positive tone. On the other hand, the results also show that better ESG performance weakens the negative relationship between the readability of sustainability reports focused on social aspects (S-Topic) and bank loan spreads. Companies with strong ESG performance can effectively reduce information asymmetry with external stakeholders, resulting in decreased informational value from the sustainability reports and diminishing the importance of readability.

The results of this study suggest that banks should consider the content of a borrowing

company's sustainability reports when conducting credit assessments. They should focus on reviewing the environmental and social aspects of these reports. Higher readability may indicate that the company is not engaging in obfuscation or hiding negative information, thus lowering the risk of default. Additionally, banks can consider the tone positivity of the report, the collateral for the loan, and the company's ESG performance. This approach allows lending banks to implement green credit policies while managing potential default risks and maximizing their ESG strategy benefits.

On the company side, when writing sustainability reports in the future, companies should use more transparent and understandable language, avoiding overly complex terms. This study recommends explaining sustainability strategies or corporate governance actions more straightforwardly to reduce banks' concerns about the report's content during credit assessments. Such efforts will enhance the information communication value of the company's sustainability report, leading to better bank loan rates and terms.

## References

- Agarwal, S., Chen, V. Y. S., & Zhang, W. N. (2016). The Information Value of Credit Rating Action Reports: A Textual Analysis. *Management Science*, 62(8), 2218-2240.
- Altman, E. (2000). Predicting Financial Distress Of Companies: Revisiting The Z-Score And Zeta. *Handbook of Research Methods and Applications in Empirical Finance*, *5*.
- Atif, M., & Ali, S. (2021). Environmental, social, and governance disclosure and default risk. Business Strategy and the Environment, 30(8), 3937-3959.
- Baier, P., Berninger, M., & Kiesel, F. (2020). Environmental, social and governance reporting in annual reports: A textual analysis. *Financial Markets, Institutions & Instruments*, 29(3), 93-118.
- Bloomfield, R. (2002). The 'Incomplete Revelation Hypothesis' and Financial Reporting. *Accounting Horizons*, *16* (3), 233-243.
- Bochkay, K., Hales, J., & Chava, S. (2020). Hyperbole or Reality? Investor Response to Extreme Language in Earnings Conference Calls. *Accounting Review*, 95(2), 31-60.
- Chiang, Y. K. (2023). Achievement Revealed Between the Lines: CSR Performance and the Readability of CSR Reports. Working paper, National Taiwan University.
- Courtis, J. K. (1998). Annual report readability variability: tests of the obfuscation hypothesis. *Accounting, Auditing & Accountability Journal, 11*(4), 459-472.
- Cui, J., Jo, H., & Na, H. (2018). Does Corporate Social Responsibility Affect Information Asymmetry? *Journal of Business Ethics*, 148(3), 549-572.
- Dale, E., & Chall, J. S. (1949). The Concept of Readability. Elementary English, 26(1), 19-26.
- Davis, A. K., Piger, J. M., & Sedor, L. M. (2012). Beyond the Numbers: Measuring the Information Content of Earnings Press Release Language. *Contemporary Accounting Research*, 29(3), 845-868.
- De Pelsmacker, P., Dens, N., & Kolomiiets, A. (2018). The impact of text valence, star rating and rated usefulness in online reviews. *International Journal of Advertising*, *37*(3), 340-359.
- Duffie, D., & Lando, D. (2001). Term Structures of Credit Spreads with Incomplete Accounting Information. *Econometrica*, 69(3), 633-664.
- Duque-Grisales, E., & Aguilera-Caracuel, J. (2021). Environmental, Social and Governance (ESG) Scores and Financial Performance of Multilatinas: Moderating Effects of Geographic International Diversification and Financial Slack. *Journal of Business Ethics*, 168(2), 315-334.
- Ertugrul, M., Lei, J., Qiu, J. P., & Wan, C. (2017). Annual Report Readability, Tone Ambiguity, and the Cost of Borrowing. *Journal of Financial and Quantitative Analysis*, *52*(2), 811-836.
- Heichl, V., & Hirsch, S. (2023). Sustainable fingerprint Using textual analysis to detect how listed EU firms report about ESG topics. *Journal of Cleaner Production*, 426, 138960.
- Henry, E. (2006). Are Investors Influenced By How Earnings Press Releases Are Written?

Journal of Business Communication, 45(4), 363-407

- Hrazdil, K., Anginer, D., Li, J., & Zhang, R. (2024). Climate Reputation and Bank Loan Contracting. *Journal of Business Ethics* 192, 875–896
- Hunjra, A. I., Jebabli, I., Thrikawala, S. S., Alawi, S. M., & Mehmood, R. (2024). How do corporate governance and corporate social responsibility affect credit risk? *Research in International Business and Finance*, 67, Article 102139.
- Iamandi, I. E., Constantin, L. G., Munteanu, S. M., & Cernat-Gruici, B. (2019). Mapping the ESG Behavior of European Companies. A Holistic Kohonen Approach. Sustainability, 11(12), Article 3276.
- Klare, G. R. (1963). *The Measurement of Readability: Useful Information for Communicators*. Iowa State University Press.
- Kuo, H. Y., Chen, S. Y., & Lai, Y. T. (2021). Investigating COVID-19 News before and after the Soft Lockdown: An Example from Taiwan. *Sustainability*, 13(20), Article 11474.
- Lee, K. H., Cin, B. C., & Lee, E. Y. (2016). Environmental Responsibility and Firm Performance: The Application of an Environmental, Social and Governance Model. *Business Strategy and the Environment*, 25(1), 40-53.
- Lo, K., Ramos, F., & Rogo, R. (2017). Earnings management and annual report readability. *Journal of Accounting & Economics*, 63(1), 1-25.
- Loughran, T., & McDonald, B. (2014). Measuring Readability in Financial Disclosures. *Journal* of Finance, 69(4), 1643-1671.
- Loughran, T., McDonald, B., & Pragidis, I. (2019). Assimilation of oil news into prices. International Review of Financial Analysis, 63, 105-118.
- Mućko, P. (2021). Sentiment analysis of CSR disclosures in annual reports of EU companies. *Procedia Computer Science*, 192, 3351-3359.
- Price, S. M., Doran, J. S., Peterson, D. R., & Bliss, B. A. (2012). Earnings conference calls and stock returns: The incremental informativeness of textual tone. *Journal of Banking & Finance*, 36(4), 992-1011.
- Raimo, N., Caragnano, A., Zito, M., Vitolla, F., & Mariani, M. (2021). Extending the benefits of ESG disclosure: The effect on the cost of debt financing. *Corporate Social Responsibility and Environmental Management*, 28(4), 1412-1421.
- Rogers, J., Van Buskirk, A., & Zechman, S. (2011). Disclosure Tone and Shareholder Litigation. *The Accounting Review*, 86(6), 2155-2183.
- Sila, I., & Cek, K. (2017). The Impact of Environmental, Social and Governance Dimensions of Corporate Social Responsibility on Economic Performance: Australian Evidence. *Procedia Computer Science*, 120, 797-804.
- Sun, Y. Q., Wang, J. J., & Huang, K. T. (2022). Does IFRS and GRI adoption impact the understandability of corporate reports by Chinese listed companies? *Accounting and Finance*, 62(2), 2879-2904.
- Suslava, K. (2021). "Stiff Business Headwinds and Uncharted Economic Waters": The Use of Euphemisms in Earnings Conference Calls. *Management Science*, 67(11), 7184-7213.

- Tsai, M.-F., & Wang, C.-J. (2017). On the risk prediction and analysis of soft information in finance reports. *European Journal of Operational Research*, *257*(1), 243-250.
- Wang, S.-M., & Ku, L.-W. (2016). ANTUSD: A Large Chinese Sentiment Dictionary. Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16), 2697–2702.
- Yu, F. (2005). Accounting transparency and the term structure of credit spreads. *Journal of Financial Economics*, 75(1), 53-84.
- Yu, K., & Garg, P. (2022). Corporate social responsibility report readability, credit ratings and cost of borrowing. *Review of Accounting and Finance*, *21*(5), 423-448.

## Appendix: Chinese and English Correspondence of the ESG Vocabulary List Proposed by Baier et al. (2020)

Word	Chinese Word	1 Chinese Word	2 Chinese Word	3 Chinese Word 4
Topic: Environment				
emission	排放	排放物	-	-
biodiversity	生物多樣	生物多樣性	生物多樣化	-
waste	廢棄物	廢料	垃圾	-
Topic: Social				
training	培訓	訓練	教育	培養
workplace	職場	工作場所	工作地點	-
poverty	貧窮	貧困	窮	-
Topic: Governance				
bribery	賄賂	受賄	收賄	賄
conformity	符合	遵從	遵守	遵照
remuneration	薪酬	薪資	酬勞	報酬

## Appendix Table 1: ESG Wordlist One-to-Many Chinese Translation Results (Excerpt)

*Note 1:* This ESG Wordlist (partial) is based on the English ESG vocabulary list proposed by Baier et al. (2020); this study compiles the remaining content.

*Note 2:* This table only includes a selection of terms. As of March 2024, the original dictionary contains 491 terms, with the number of terms in the Environmental, Social, and Governance categories being 64, 151, and 276, respectively.

Voor		ESG_SEGMENT							
Ieal	(1) E-Topic	(2) S-Topic	(3) G-Topic	(4) ESG-Topic	(5) Total				
2014	76	125	86	16	271				
2015	94	149	113	21	304				
2016	114	187	131	23	374				
2017	144	205	170	37	407				
2018	158	229	192	54	427				
2019	202	261	223	71	476				
2020	248	281	263	99	523				
2021	430	323	385	178	648				
2022	671	377	574	247	843				
Total	2,137	2,137	2,137	746	4,273				

 Table 1. Classification Results of Sustainability Report Samples into

Four	Main	Themes
------	------	--------

Note: Sustainability report texts primarily focused on the Environmental aspect (E-Topic) are classified as such when the proportion of Environmental (E) keywords in the company's text exceeds the median of all samples. Similarly, texts focused on the Social aspect (S-Topic) are classified when the proportion of Social (S) keywords is greater than the median, and those focused on Governance (G-Topic) are classified when the proportion of Governance (G) keywords exceeds the median. Texts defined as E-Topic, S-Topic, and G-Topic simultaneously are categorized as Integrated ESG Three Themes (ESG-Topic) texts.

## Table 2. Classification Results of Positive Emotional Tone in

		ESG_SEGM	ENT	
-	E-Topic	S-Topic	G-Topic	ESG-Topic
TONE_SCORE				
Mean	0.434	0.481	0.493	0.442
TONE_SEGMENT				
High	845	986	1043	299
%	39.5%	46.1%	48.8%	40.1%
Low	1,292	1,151	1,094	447
%	60.5%	53.9%	51.2%	59.9%
Total	2,137	2,137	2,137	746

## Sustainability Report Samples

Variables	Obs	Mean	Std Dev	Min	Median	Max
LOAN_SPREAD	13,769	0.472	0.594	-0.217	0.305	1.961
COMPLEX_WORDS_PERCENT	13,769	0.339	0.023	0.268	0.339	0.426
FIRM_SIZE	13,769	17.444	1.554	12.141	17.444	22.087
LEVERAGE	13,769	0.190	0.139	0.000	0.168	0.964
МВ	13,769	0.654	0.609	0.013	0.480	10.091
PROFITABILITY	13,769	0.088	0.063	-0.405	0.079	0.514
TANGIBILITY	13,769	0.306	0.182	0.000	0.290	0.924
BANKRUPT	13,769	2.303	1.617	-1.644	1.959	22.042
TCRI	13,769	529.341	222.376	-842	552	1,015
ESGP	13,769	60.334	7.188	34.810	59.950	81.840
TONE_SCORE	13,769	0.123	0.013	0.083	0.123	0.171
LOAN_SIZE	13,769	12.158	1.646	0.000	12.206	18.299
MATURITY	13,769	1.888	2.491	0.010	1.000	24.220
SECURITY	13,769	0.239	0.427	0.000	0.000	1.000

Table 3. Descriptive Statistics for the Full Sample

Note: *FIRM<sub>SIZ</sub>* is defined as the natural logarithm of total assets, *LEVERAGE* is defined as the ratio of long-term debt to total assets, *MB* is defined as the market value of common stock at the end of the period divided by total equity, *PROFITABILITY* is calculated as EBITDA divided by total assets, *TANGIBILITY* is measured as the ratio of long-term tangible assets to the book total of Property, Plant, and Equipment (PPE). *BANKRUPT* is assessed using the Z-score defined by Altman (2000), where a score closer to 0 indicates a potential bankruptcy crisis, while a score closer to 3 indicates smore muscular financial health, *TCRI* refers to the total score of the Taiwan Corporate Credit Risk Indicator developed by TEJ, *ESGP* is the total score of the sustainability report, with positive values indicating a positive tone and negative values indicating a negative tone, Regarding loan characteristics: *LOAN\_SIZE* is defined as the natural logarithm of the total loan amount, *MATURITY* refers to the loan term, measured in years, *SECURITY* is a dummy variable where secured loans are coded as 1 and unsecured loans as 0.

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
LOAN_SPREAD (1)													
COMPLEX_WORDS_PERCENT (2)	-0.104*												
FIRM_SIZE (3)	-0.003	0.312*											
LEVERAGE (4)	-0.017	0.178*	0.254*										
<i>MB</i> (5)	-0.168*	-0.012	-0.265*	-0.226*									
PROFITABILITY (6)	-0.226*	0.091*	0.010	-0.110*	0.598*								
TANGIBILITY (7)	-0.079*	0.042*	-0.093*	0.378*	0.066*	0.206*							
BANKRUPT (8)	-0.141*	-0.013	-0.194*	-0.398*	0.833*	0.622*	-0.053*						
<i>TCRI</i> (9)	-0.206*	0.116*	0.266*	-0.256*	0.325*	0.554*	-0.082*	0.421*					
<i>ESGP</i> (10)	-0.093*	0.090*	0.139*	0.119*	0.110*	0.161*	0.108*	0.119*	0.118*				
TONE_SCORE (11)	-0.024*	0.187*	0.074*	-0.001	0.059*	0.093*	-0.163*	0.063*	0.085*	-0.001			
LOAN_SIZE (12)	-0.126*	0.192*	0.570*	0.058*	-0.054*	0.104*	-0.066*	-0.028*	0.227*	0.125*	0.086*		
MATURITY (13)	0.035*	-0.043*	0.020*	0.189*	-0.023*	0.041*	0.125*	-0.062*	-0.030*	0.016	-0.005	0.126*	
SECURITY (14)	0.221*	-0.130*	-0.131*	0.021*	-0.104*	-0.202*	-0.091*	-0.157*	-0.213*	-0.175*	-0.060*	-0.025*	0.295*

## **Table 4. Correlation Coefficient Matrix**

Note 1: The correlation coefficient matrix is calculated using Pearson correlation coefficients.

Note 2: This table uses annual company data, with 13,769 observations, excluding any samples with missing values for any variable. Due to a company potentially having multiple loan records in a single fiscal year, the number of observations here is higher than that of the sustainability report samples.

Note 3: Please refer to Table 3 for definitions of each variable.

Note 4: An asterisk (\*) indicates significance at the 5% level in a two-tailed test.

Variables		Dependent	Variable: LOA	AN_SPREAD	
variables	(1) Total	(2) E-Topic	(3) S-Topic	(4) G-Topic	(5) ESG-Topic
COMPLEX	1.409***	1.965***	2.158***	0.822	-2.206
WORDS PERCENT	(3.35)	(2.92)	(3.01)	(1.07)	(-1.33)
FIRM_SIZE	-0.087***	-0.119*	-0.116**	-0.332***	-0.314**
_	(-2.75)	(-1.70)	(-2.35)	(-5.43)	(-2.32)
LEVERAGE	-0.270***	-0.356*	0.275	0.436**	0.308
	(-2.63)	(-1.94)	(1.46)	(2.12)	(0.6)
MB	-0.039	-0.159***	0.006	-0.020	-0.013
	(-1.46)	(-3.74)	(0.13)	(-0.50)	(-0.14)
PROFITABILITY	-0.636***	-0.192	$-1.245^{***}$	$-0.595^{***}$	-0.691
	(-4.73)	(-0.65)	(-6.27)	(-3.11)	(-1.05)
TANGIBILITY	-0.790***	$-1.480^{***}$	-1.103***	$-0.657^{***}$	-0.359
	(-6.85)	(-7.23)	(-6.56)	(-2.83)	(-0.88)
BANKRUPT	-0.027***	-0.012	-0.023	-0.011	-0.012
	(-2.77)	(-0.84)	(-1.18)	(-0.64)	(-0.31)
TCRI	0.000	-0.000**	0.000	0.000*	0.000
	(1.25)	(-2.16)	(0.38)	(1.66)	(1.59)
ESGP	-0.005***	-0.014***	-0.009***	-0.002	-0.017***
	(-3.59)	(-6.68)	(-4.14)	(-0.60)	(-2.95)
TONE_SCORE	-2.078***	-0.119	$-3.146^{***}$	1.991	-1.749
	(-2.92)	(-0.10)	(-2.93)	(1.56)	(-0.56)
LOAN_SIZE	-0.032***	$-0.046^{***}$	-0.014**	-0.044***	0.020
	(-7.75)	(-7.47)	(-2.54)	(-6.62)	(1.53)
MATURITY	0.006***	-0.001	0.003	0.006**	-0.001
	(2.75)	(-0.50)	(1.06)	(1.97)	(-0.30)
SECURITY	0.025**	0.047***	0.010	0.059***	0.017
	(2.15)	(2.60)	(0.60)	(3.08)	(0.45)
Fixed Effect (FE)	Year, Firm,	Year, Firm,	Year, Firm,	Year, Firm,	Year, Firm,
	Bank	Bank	Bank	Bank	Bank
Adjusted R <sup>2</sup>	0.567	0.539	0.552	0.609	0.647
Observations	13,710	7,001	7,790	4,805	1,532
Coefficient Difference	Test				
(3) - (2)			0.48		
(3) - (4)			11.33***		
(3) – (5)			10.88***		

Table 5. The Relationship Between Sustainability Report Readability and Bank LoanSpreads for Each Subsample

Note 1: \*\*\*, \*\*, \* indicate statistical significance at the 1%, 5%, and 10% levels, respectively, with the t-statistic in parentheses.

Note 2: The number of observations in the primary model is lower than the original sample size due to removing individual observations that could not provide within-group variation after controlling for fixed effects of year, company, and lending bank.

Note 3: For definitions of variables, please refer to Table 3.

Variables	First S COMPLEX_WOR	Stage RDS_PERCENT	Second Stage LOAN_SPREAD		
	(1) E-Topic	(2) S-Topic	(3) E-Topic	(4) S-Topic	
AvgG_ComplexW	0.885***	0.956***			
	(80.42)	(121.95)			
P_ComplexW			2.387***	1.790**	
			(2.88)	(2.24)	
Controls	Yes	Yes	Yes	Yes	
Fixed Effect (FE)	Year, Firm	Year, Firm	Year, Firm	Year, Firm	
Adjusted R <sup>2</sup>	0.968	0.984	0.528	0.543	
Observations	7,040	7,830	7,040	7,830	
F	980.15	3998.70	-	-	

# Table 6. Instrumental Variable Regression Results – Samples Focused on Environmental (E-Topic) and Social (S-Topic) Aspects

Note 1: \*\*\*, \*\*, \* indicate statistical significance at the 1%, 5%, and 10% levels, respectively, with the t-statistic in parentheses.

Note 2: Control Variables: All control variables are consistent with the primary model.

Note 3: Instrumental Variable: This study uses the average proportion of complex words ( $AvgG\_ComplexW$ ) in sustainability reports from companies within the same group for the same year as the instrumental variable, with group composition data sourced from TEJ.

Note 4:Predicted Value: *P* ComplexW represents the predicted value from the first-stage regression.

	<u>First Stage</u>	<u>Se</u>	cond Stage		
Variables	Dependent Variable:	Dependent Variable:			
Variables	ESG_REPORT	LOA	AN_SPREAD		
	(1)	(2) Full Sample	(3) E-Topic	(4) S-Topic	
COMPLEX _WORD_PERCENT		1.192 <sup>***</sup> (2.86)	1.816 <sup>***</sup> (2.72)	2.205 <sup>***</sup> (3.14)	
FIRM_SIZE	0.580*** (65.99)	-0.070* (-1.93)	-0.104 $(-1.41)$	$-0.160^{***}$ (-2.98)	
LEVERAGE	$-0.825^{***}$ (-10.12)	$-0.293^{***}$ (-2.79)	$-0.370^{**}$ (-2.01)	$0.334^{*}$ (-1.74)	
МВ	0.049*** (-3.23)	-0.036 $(-1.31)$	$-0.147^{***}$ (-3.32)	-0.014 $(-0.27)$	
PROFITABILITY	$-1.882^{***}$ (-12.56)	$-0.622^{***}$ (-4.43)	-0.27 $(-0.90)$	$-0.976^{***}$ (-4.83)	
TANGIBILITY	1.019*** (20.62)	$-0.752^{***}$ (-6.29)	$-1.401^{***}$ (-6.44)	$-1.267^{***}$ (-7.10)	
BANKRUPT	-0.003 (-1.54)	$-0.028^{***}$ (-2.90)	-0.015 (-1.02)	-0.017 $(-0.88)$	
TCRI	0.001*** (17.94)	0.000* (1.68)	0.000 (1.50)	0.000 (0.86)	
ESGP	0.063*** (52.26)	-0.004 $(-1.60)$	$-0.012^{***}$ (-4.00)	$-0.014^{***}$ (-4.35)	
LOAN_SIZE	$-0.047^{***}$ ( $-7.42$ )	$-0.033^{***}$ (-7.72)	$-0.048^{***}$ (-7.48)	-0.010* (-1.67)	
MATURITY	$-0.010^{***}$ (-2.88)	0.005** (2.44)	-0.002 $(-0.60)$	0.003 (1.14)	
SECURITY	0.063*** (3.36)	0.027** (2.33)	0.049*** (2.67)	0.006 (0.39)	
TONE_SCORE		$-1.946^{***}$ (-2.73)	-0.041 (-0.03)	-3.149*** (-2.95)	
LAMBDA		0.062 (1.10)	0.084 (0.96)	$-0.203^{**}$ (-2.54)	
Fixed Effect (FE)	Year	Year, Firm, Bank	Year, Firm, Bank	Year, Firm, Bank	
Adjusted/Pseudo R <sup>2</sup>	0.408	0.566	0.539	0.551	
Observations	42,764	13,717	7,001	7,797	
F	21978.92	63.17	36.73	33.41	

## Table 7. Heckman Two-Stage Sample Selection Model Regression Results

Note 1: \*\*\*, \*\*, \* indicate statistical significance at the 1%, 5%, and 10% levels, respectively, with the t-statistic in parentheses.

**Note 2:** Among all sample companies with loan data, if the company prepared a sustainability report in that year, ESG\_REPORT is coded as 1; otherwise, it is coded as 0.

Variables	Dependent Variable: LOAN_SPREAD					
		S-Topic				
TONE_SEGMENT	(1) High	(2) Low	(3) High	(4) Low		
COMPLEX _WORD_PERCENT	1.239 (1.10)	0.652 (0.63)	6.549*** (5.44)	0.830 (0.82)		
Controls	Yes	Yes	Yes	Yes		
Fixed Effect (FE)	Year, Firm, Bank	Year, Firm, Bank	Year, Firm, Bank	Year, Firm, Bank		
Adjusted R <sup>2</sup>	0.572	0.555	0.566	0.583		
Observations	2,946	4,034	3,754	4,021		

# Table 8. Regression Results of the Moderation Model: Using Positive Tone Text as the Moderating Variable

Note 1: \*\*\*, \*\*, \* indicate statistical significance at the 1%, 5%, and 10% levels, respectively, with the t-statistic in parentheses.

Note 2:Control Variables: All control variables are consistent with the primary model.

Variables	Dependent Variable: LOAN_SPREAD			
	(1) <i>E-Topic</i>	(2) S-Topic		
COMPLEX_WORDS_PERCENT	-6.099	12.382**		
	(-1.52)	(2.41)		
ESGP	-0.058***	0.046*		
	(-2.68)	(1.65)		
COMPLEX_WORDS_PERCENT × ESGP	0.126**	-0.161*		
	(2.01)	(-1.96)		
Controls	Yes	Yes		
Fixed Effect (FE)	Year, Firm, Bank	Year, Firm, Bank		
Adjusted R <sup>2</sup>	0.539	0.553		
Observations	7,001	7,790		
NT / 1 *** ** * 1. / / / 1. 1 /	1 10/ 50/ 1 100/ 1 1			

## Table 9. Regression Results of the Moderation Model: Using ESG Performance as the Moderating Variable

Note 1: \*\*\*, \*\*, \* indicate statistical significance at the 1%, 5%, and 10% levels, respectively, with the t-statistic in parentheses.

Note 2:Control Variables: All control variables are consistent with the primary model.