Investor's ESG Tendency Probed by Pre-trained Transformers

Li, Chao and Keeley, Alexander Ryota and Takeda, Shutaro and Seki, Daikichi and Managi, Shunsuke

Kyushu University

November 2024

Investor's ESG Tendency Probed by Pre-trained Transformers

Abstract

Due to climate change and social issues, environmental, social, and governance (ESG) solutions receive increased attention and emphasis. Being influential market leaders, investors wield significant power to persuade companies to prioritize ESG considerations. However, investors' preferences for specific ESG topics and changing trends in those preferences remain elusive. Here, we build a group of large language models with 128 million parameters, named classification pre-trained transformers (CPTs), to extract the investors' tendencies toward 13 ESG-related topics from their annual reports. Assisted by the CPT models with approximately 95% cross-validation accuracy, more than 3,000 annual reports released by globally 350 top investors during 2010-2021 are analyzed. Results indicate that although the investors show the strongest tendency toward the economic aspect in their annual reports, the emphasis is gradually reducing and shifting to environmental and social aspects. Nonfinancial investors like corporation and holding company investors prioritize environmental and social factors, whereas financial investors pay the most attention to governance risk. There are differences in patterns at the country level, for instance, Japan's investors show a greater focus on environmental and social factors than other major countries. Our findings suggest that investors are increasingly valuing sustainability in their decision-making. Different investor businesses may encounter unique ESG challenges, necessitating individualized strategies. Companies should improve their ESG

disclosures, which are increasingly focused on environmental and social issues, to meet investor expectations and bolster transparency.

Keywords:

ESG; Investor; Natural Language Processing; Pre-trained Transformer; Data Mining; Machine Learning

Introduction

Recently, the urge to advance corporate sustainability and promote corporate social responsibility (CSR) has garnered considerable attention, especially in light of the emergence of various environmental issues and public health disasters (G20 Green Finance Study Group, 2017; Global Sustainable Investment Alliance, 2017; IPCC, 2022). Environmental, social, and governance (ESG) is a novel, all-encompassing concept proposed by the United Nation's Principles for Responsible Investment report that covers the environmental, social, and economic performance of corporations (Rahman, Zahid, & Al-Faryan, 2023; Zhou, Liu, & Luo, 2022). Various environmental and social issues, such as climate change (IPCC, 2022), air pollution (Lelieveld, Evans, Fnais, Giannadaki, & Pozzer, 2015), water pollution (Wu, Palm-Forster, & Messer, 2021), violation of human rights (Schrempf-Stirling & Wettstein, 2017), and poor work environments (Furman et al., 2019), compel societies, governments, corporations, and the general public to make adjustments. Rather than focusing solely on short-term financial benefits, investors, regulators, lawmakers, and customers are increasingly requiring corporations to formulate strategies and planning that consider and balance environmental, social, and economic performance (Alshehhi, Nobanee, & Khare, 2018; Zhou et al., 2022). If corporations do not adhere to ESG and fulfill their social responsibility following the signing of key regulations, they may lose market share in the short term and be eliminated in the long run (Gustafsson, Schilling-Vacaflor, & Lenschow, 2023; Schilling-Vacaflor, 2021). Therefore, societies, corporations, investors, and even individuals must take ESG into account.

Investors' attitude toward ESG dramatically affects corporate actions because companies require sufficient investment from investors (Barko, Cremers, & Renneboog, 2022; Diener, 2023; Espahbodi, Espahbodi, Juma, & Westbrook, 2019; Raut, Shastri, Mishra, & Tiwari, 2023). Investor engagement strategies have a significant impact on impact-oriented investment (Diener, 2023). Furthermore, investors' perceptions of the relevance and reliability of ESG information should be improved, as it has a long-term mediating impact on the investors' investment allocation and price assessment (Espahbodi et al., 2019). Previous studies suggest that investors disclose their financial incentives for ESG (Espahbodi et al., 2019; Raut et al., 2023). However, to our knowledge, no global unified regulations exist to claim the ESG tendency of investors standardly. In other words, it is difficult to assess investors' attitudes toward ESG using a globally recognized benchmark. Instead of relying solely on disclosed numbers (Heichl & Hirsch, 2023; Loughran & McDonald, 2011, 2016), the language used by corporations can reveal their attitude toward ESG. The primary goal of this study is to address the following issues: (1) the general tendency of investors toward various ESG-related topics; (2) changing trend of investors' tendency; (3) whether investors belonging to different categories show a relatively different tendency composition; and (4) whether differences exist in the structure of investors' tendency in various major countries.

In this study, we consider and investigate 13 ESG-related topics. Consistent with previous research, they cover the entire ESG, which can be divided into three studies: environmental aspect, social aspect, and economic aspect (Feng, Zhu, & Lai, 2017; Gillan, Koch, & Starks, 2021; Alexander R Keeley et al., 2022; Xie, Nozawa, Yagi, Fujii, & Managi, 2019). In particular, the environmental aspect comprises four topics: air pollution, greenhouse gases, water consumption, and mining consumption; the social aspect comprises six topics: safety and health, human rights, work environment, community, domestic job creation, and domestic reflux rate; and the economic aspect comprises governance risk, production cost, and economic ripple effect. Prior to ESG considerations, corporations are viewed as profit-driven economic actors, while governments, as political roles and supervisors, are responsible for societal development (Sundaram & Inkpen, 2004). However, as a result of rapid globalization, companies, particularly multinational corporations, should assume greater social responsibility and increase their transparency in environmental and social matters (Rahman et al., 2023; Tang, Xiong, & Peng, 2024). Environmentally, the factory's air pollution has a direct correlation with local public health (Dedoussi, Eastham, Monier, & Barrett, 2020; Lelieveld et al., 2015). Climate change issues resulting from greenhouse gas emissions are noticeable (Burrell, Evans, & De Kauwe, 2020; IPCC, 2022; Ortiz-Bobea, Ault, Carrillo, Chambers, & Lobell, 2021). Water scarcity and mining-caused soil contamination impact human daily life (Cheng, Hu, & Zhao, 2009; Li, Ma, van der Kuijp, Yuan, & Huang, 2014). The majority of these environmental changes are attributable to corporate activities. In social aspects, the impact of corporations on the local community is gradually gaining more attention, including community (Ismail, 2009), occupational health and safety (Montero, Araque, & Rey, 2009), human rights (Giuliani, 2016), and local impacts (Panthong & Taecharungroj, 2021). According to the Taskforce on Nature-related Financial Disclosures (TNDF) (https://tnfd.global/), financial institutions now require corporations to disclose nature-related information. Standards, such as social accountability 8000 (SA8000, https://sa-intl.org/programs/sa8000/) and International Labour Organization (ILO) conventions, list a number of specific CSR-related items, which our topics refer to. Obviously, laws, and regulations, such as the French Duty of Vigilance law and the Non-Financial Reporting Directive, are essential references. The

13 indexes listed above were selected based on laws, regulations, organization guidelines, standards, and scholarly research.

ESG scores from several major ESG rating agencies, including Refinitiv's ESG score (LSEG, 2023), Kinder Lydenberg, and Domini (KLD) (Sharfman, 1996), Sustainalytics (Harrison, Yu, & Zhang, 2023), Moody's ESG (Kiesel & Lücke, 2019), S&P Global (Gyönyörová, Stachoň, & Stašek, 2023), and MSCI (Sabbaghi, 2022), have significant implications. However, each agency covers a wide range of ESG-related topics and metrics, and their ESG scores differ significantly (Berg, Kölbel, & Rigobon, 2022; Senadheera et al., 2021; Wan, Dawod, Chanaim, & Ramasamy, 2023). The divergence of rating methods stems from different category taxonomies. Refinitiv's ESG score, the most widely used ESG database among them, has the most comprehensive categories and is based on three pillar scores: the E score, S score, and G score (Berg et al., 2022; LSEG, 2023). The E score has three categories: resource use, emissions, and innovation; the S score contains four categories: workforce, human rights, community, and product responsibility; finally, the G score focuses on management, shareholders, and CRS strategy. For instance, the MSCI's environmental pillar addresses four major issues: climate change, natural capital, pollution & waste, and environmental opportunities; the social pillar addresses human capital, product liability, stakeholder opposition, and social opportunities; and the governance pillar addresses corporate governance and corporate behavior (MSCI, 2024). In the MSCI ESG score framework, each issue contains several topics; for example, in the MSCI ESG score framework, carbon emissions fall under climate change, natural capital covers raw material sourcing, and health & safety fall under human capital (MSCI, 2023a). These frames from the widely used ESG database are another resource for determining the 13 potential topics. Importantly, the names of these topics differ between databases, and some terms are uncommon in the annual reports, so we have literally simplified and summarized these topics.

The listed companies publish their annual reports in order to communicate with their stakeholders, increase transparency, and accountability, comply with legal, and regulatory requirements, maintain investor relations, disclose performance evaluation and planning, and improve their market reputation and brand image (Ramzan, Amin, & Abbas, 2021; Stanton & Stanton, 2002). The annual report of a company serves a much greater purpose: it implicitly reflects the company's stance on a particular subject. Nonfinancial reports, such as sustainability reports and CRS reports, provide more detailed disclosure than annual reports and have become popular in recent years (Heichl & Hirsch, 2023; Skouloudis, Evangelinos, & Kourmousis, 2010). Previous studies have used these reports to analyze company insights on specific topics (Goloshchapova, Poon, Pritchard, & Reed, 2019; Landrum & Ohsowski, 2018). However, nonfinancial reports differ significantly across countries and company sizes (Dissanayake, Tilt, & Xydias-Lobo, 2016; Skouloudis et al., 2010). Furthermore, some investors have yet to release their sustainability or CSR reports, despite an increasing number of companies doing so (Christensen, Hail, & Leuz, 2021; Opferkuch, Caeiro, Salomone, & Ramos, 2021). The materials chosen for this study are annual reports, that are routine, stable, and long-standing. Annual reports are concise and informative, so the amount of representation of a topic within them often reflects its importance. Therefore, accurately determining the proportion of ESG-related content may provide insight into companies' focus on this issue. Furthermore, annual reports frequently include discussion regarding CSR, the central issue of ESG (Anas, Abdul Rashid, & Annuar, 2015; Chan, Watson, & Woodliff, 2014; Chijoke-Mgbame, Mgbame, Akintoye, & Ohalehi, 2019). Therefore, we use annual reports of investors as analysis materials in this study.

Textual analysis has been used in the past to estimate the ESG tendency of corporations by calculating the ratio of ESG-related words in their annual reports (Baier, Berninger, & Kiesel, 2020; Caglio, Melloni, & Perego, 2020). However, these studies have either stringent word count requirements or a restricted number of reports to be analyzed. As generative pre-trained transformers (GPT), such as ChatGPT, become more widely used, and recognized (Radford, Narasimhan, Salimans, & Sutskever, 2018), large-scale artificial intelligence (AI) natural language processing (NLP) models are gradually expanding across a variety of industries. FinBERT (Huang, Wang, & Yang, 2023), KoBERT (Bang, Ryu, & Yu, 2023), and ESGBERT (Mehra, Louka, & Zhang, 2022) are some of the advanced NLP models developed for ESG analysis. These models are based on fine-tuning the Bidirectional Encoder Representations from Transformers (BERT) model, which may explain why they have lower accuracy than our models. Large language models outperform other machine learning methods, like naïve Bayes, support vector machine, and random forest, showing the potential and broad prospects of large language models in ESG-related and financial analyses (Huang et al., 2023). To conduct a comprehensive and effective analysis of unstructured text data, such as annual reports, we construct a set of machine learning models, CPTs, following the transfer learning technique. To make our model more suitable for our task, we redesigned the network architecture. The novel model is designated CPT because it is a classification model for “C,” it transplants layers from a pre-trained model for “P,” and its core component is the transformer block for “T.” Transformer block layer could help the models focus on the keywords and their pattern to make accurate predictions (Vaswani et al., 2017). To reduce the number of parameters that must be trained, a portion of Google’s BERT model (Devlin, Chang, Lee, & Toutanova, 2018) is transplanted. We fragment each investor company’s annual report into 60-word chunks.

Then, we assess the relevance of each fragment of an annual report to a particular ESG-related topic. Finally, we aggregate all relatedness scores of all fragments in the annual report into a single value ranging from 0 to 1, which represents the trend of the annual report on the ESG-related topic that was analyzed. We train 13 CPTs, each for a single ESG-related topic.

This study contributes to the literature in several aspects. First, this study examines the tendencies of global investors on several topics, as well as their temporal variations, using annual reports. Several previous studies show a significant shift in investor preferences from purely economic considerations to a greater focus on ESG issues such as carbon emissions and social responsibility (Bolton & Kacperczyk, 2021; Chatzitheodorou, Skouloudis, Evangelinos, & Nikolaou, 2019; Krüger, 2015). However, these studies typically describe a broad trend shift without delving into specific annual fluctuations. Furthermore, prior research primarily examined the circumstances within specific countries, such as investor perceptions in South Korea (Park & Jang, 2021), pillar importance in India (Sood, Pathak, Jain, & Gupta, 2023), among many others (Tang et al., 2024). Second, we investigated and compared the disparities among different sub-style industrial investors classified on the basis of the Global Industry Classification Standard (GICS) (MSCI, 2023b). The findings show that investor preferences are somewhat consistent with their sub-styles. According to Rojo-Suárez and Alonso-Conde (2024), the majority of research focuses on the examination of specific sectors or stock categories, for example, Hong, Li, and Xu (2019) and Blitz and Fabozzi (2017) rather than providing a comprehensive analysis of all industries across the economy. Third, we examine the differences between investors in several major countries using the same benchmark. Cross-country comparisons are uncommon due to a scarcity of large-scale studies on investor inclinations toward specific ESG topics. Fourth, a novel textual analysis method that can objectively and efficiently analyze annual reports is another major contribution. In addition to extracting quantitative data from narrative data, there is a noticeable trend of textual analyses on annual reports being used more frequently in accounting and finance to address diverse inquiries (Baier et al., 2020; Lokuwaduge & Heenetigala, 2017; Loughran, McDonald, & Yun, 2009). However, those studies either require experienced researchers to manually classify the sentences (Giles & Murphy, 2016; Tilling & Tilt, 2010), or require the creation of comprehensive dictionaries (Baier et al., 2020; Bodnaruk, Loughran, & McDonald, 2015). Our CPT models combine the strengths of the two methods mentioned above to improve textual analysis by considering the semantic context of each phrase and reducing human reliance.

From 2010 to 2021, this study develops high-accuracy CPT models to analyze more than 3,000 annual reports from 350 top investors globally with the highest ESG investment ownership. The tendencies on 13 ESG-related topics are extracted from reports containing more than 102 million valid words. These investors' average tendencies are summarized. Additionally, the annual changes for each ESG-related subject are investigated. We examine further the distinction between investors with various primary businesses and the temporal variation between the various types.

Methods

Materials

Annual reports of investors

This study aims to develop models for analyzing the ESG tendency of major capital market investors. A previous study determined the top 350 investors with the loudest voice on ESG in the 2020 financial year (Keeley, Li, Takeda, Gloria, & Managi, 2022). Specifically, we focus on investors who have the most investment in companies with high ESG scores. Alexander R. Keeley et al. (2022) obtained the investors' investment portfolio and the ESG scores of the invested companies. Then, the ranking of investors is computed by the accumulated values of the products of the ESG scores and the total investment amount of all invested companies. For example, if an investor invests one million dollars in ten companies each having an ESG score of 0.8, the investor will own eight million ESG shares. The computation is more complicated in practice; refer Alexander R. Keeley et al. (2022) for more information. In this way, two factors affect an investor's rank: first, the amount of capital under management, and second, the ESG performance of invested companies. According to our analysis of the investor list (**Supplementary Materials Table S1**: Investor List), the scale of capital has a greater impact. We explore the 13 ESG topics in greater detail using the annual reports of these investors in our analysis. The primary method is the textual analysis of the annual report's language. These annual reports are generally available to the public and listed on the investors' websites. While building the annual report dataset, we prioritize user-friendly annual reports and integrated reports. If neither of these two types of reports is available, annual reports on Form 10-K or Form 20-F are also acceptable. According to the list of the top 350 investors, their annual reports from 2010 to 2021 are downloaded. All of these annual reports are in English versions. We obtain 3,217 PDF files representing annual reports from the internet.

Natural Language Processing (NLP)

Data preprocessing is required because PDF files cannot be utilized directly for analysis. First, we must extract text content from PDF documents. The extracted text retains only uppercase and lowercase letters and numbers, and uppercase letters are converted to lowercase letters. The retained words are simplified by returning verbs to their original forms and singularizing plural nouns. This process will reduce the variety of words in a text's context. A lower word diversity is associated with a shorter tokenization time and less computer resource usage. To shorten the context, we remove stop words, such as "a," "an," "the," "to," and "and." The context is tokenized with a pre-trained tokenizer, the BERT-base-uncased tokenizer (Devlin et al., 2018). Finally, the raw data are converted to the BERT-base-uncased tokenizer's word IDs. The BERT-base-uncased tokenizer is not case-sensitive; for instance, there is no difference between "Apple" and "apple" during the encoding process. Furthermore, to reduce dictionary size, the BERT-base-uncased tokenizer employs subword tokenization technology. For example, "going" would be separated into "go" and "##ing" and then converted into two numbers respectively. The BERT word IDs are unique integers corresponding to a single subword. Using the BERT tokenizer is required because the BERT embedding layer is being transplanted to improve the performance of our model. Since some PDF files are unreadable by machines, the BERT tokenizer successfully encodes 2,520 files in total. 102,750,289 word IDs are extracted from these 2,520 files.

The input data for our NLP model are lists of BERT word IDs with fixed lengths, also known as fragments. In this study, the length was set at 60 words. We need to clarify that the "word" here actually refers to "token." Cross-validation indicates that there is no significant difference in model performance between 60-word and 80-word lengths; however, using a fragment with 80 words requires more computer resources and time. The model trained with 60-word fragment data could perform better than the model trained with 40-word fragment data. Therefore, the 60-word analysis length is reasonable. The dataset of fragment lists is referred to as the NLP model's input data.

The input dataset is generated as follows: first, the cursor points to the first words of an encoded annual report; second, the first 60 words from the cursor location are sliced into the first fragment; third, the two-word IDs adjacent to the first 60 words are considered raw output data; and finally, the cursor advances one word and repeats the previous three steps until the raw output data reach the last word of the annual report. We generate the input dataset report by report, and the first 60 words and last 60 words of each report cannot be put into 60 fragments; therefore, the total data size is less than 102,750,289, or approximately 102 million. The raw output data is a two-word phrase. The real output data is a binary value, with 1 representing yes, and 0 representing no. On the basis of the judgments, the raw output data should be transformed into the output data. The determination is whether the two-word phrase is on the list of predefined phrases. If the phrases in the raw output data are in the predefined phrase list, the output data are labeled with a 1; otherwise, they are labeled with a 0. In the following section, we explain the predefined word list in detail.

ESG-related topic phrases (ERTPs)

This study is interested in ESG-related topics, including human rights, governance risk, greenhouse gases, safety and health, mining consumption, community, domestic job creation, domestic reflux rate, production cost, water consumption, air pollution, economic ripple effect, and work environment. Using two BERT word IDs to encode the ERTPs reduces the amount of work required to create a training dataset. Several phrases contain three or more words, so we have shortened and simplified them. "human right," "governance risk," "greenhouse gas," "safety health," "mining resource," "local community," "domestic job," "domestic reflux," "production cost," "water consumption," "water consumption," "air pollution," "ripple effect," and "work environment" are the initial ERTPs. Since the tendency of ESG-related topics will be analyzed separately, the output data will be labeled 13 times, such as whether they are consistent with human rights or governance risk. To mark the raw output data, we determine if the phrase in the raw output phrase data is identical to an ERTP. Even within a 102 million-record dataset, strict consistency is uncommon. Therefore, each ERTP phrase list must be expanded. The literal synonyms of ERTPs are evaluated first. If the phrase is a synonym of one ERTP, it will be added to the phrase list of that ERTP. It must be emphasized that the ERTPs' synonyms are identical or similar phrases. For example, "carbon dioxide" and "greenhouse gas" are considered synonymous in the majority of scenarios, so "carbon dioxide" should be added to the ERTP phrase list for "greenhouse gas." Without enough positive data, valid large-scale NLP model training is difficult. Nonetheless, the number of literal synonyms must be limited. Risky is the expansion of ERTP synonym lists without careful consideration.

In the study, a semi-automatic program dubbed "ESG synonyms searcher" (ESS) was developed to effectively and efficiently search for synonyms. In essence, our method is a commonly used NLP learning technology known as active learning. Active learning is semi-supervised. This method uses iterative training and prediction to continuously update the data labeling criteria (Roh, Heo, & Whang, 2021; Schröder & Niekler, 2020). Manual data labeling is much more expensive and time-consuming than active learning. The search processing is as follows: first, we build a training dataset for a CPT model and train the CPT model; second, we randomly select some input data that are labeled as 0 and use the trained CPT model to predict that dataset; third, we analyze the predicted result to find which phrase is always mislabeled as positive and take the top 30 as the potential ERTP synonyms; fourth, we calculate the cosine similarity between phrases in the synonym lists of a certain ERTP and phrases obtained in the third step; fifth, we artificially check whether the phrase with over 60% similarity is a synonym of the ERTP; lastly, we repeat the first six steps until the synonym lists of ERTP is large enough. Here is a list of information regarding the ESS. In the first step of constructing the training dataset, all positive-labeled input data is selected based on the current ERTP synonym phrase list. Then, 20 random samples of input data labeled 0 are taken. We combine these datasets and assign weights to each record to balance the accurate and inaccurate data ratio. In the second step, we train the CPT model with 50 epochs and set the early stop with 20 epochs of patience, based on the loss of the cross-validation dataset.

Procedures

Basic assumptions

Here, to precisely analyze the annual reports, we propose two assumptions: first, companies care more about the topic they frequently mention in annual reports than topics that do not appear at all (**Assumption I**), which is widely adopted in previous studies (Baier et al., 2020; Bodnaruk et al., 2015); second, some parts of a sentence must be highly related to another meaningful part of the same sentence's interest (**Assumption II**). In other words, the purposeful parts can be anticipated and deduced from their context. For instance, if a sentence fragment reads, "air pollution causes some cancer, so we plan to reduce (UNKNOWN)," most people would assume that (UNKNOWN) refers to a phrase associated with air pollution. Then, we generalize **Assumption II,** which states that a specific phrase can be inferred from the words that precede it. The GPT model produces a sentence that conforms to the same logic (Radford et al., 2018; Radford et al., 2019). This generalization holds true for certain types of texts, especially company annual reports, which are typically succinct, coherent, and easily understood without the use of metaphor or irony.

Based on **Assumption II**, after reading a few words, it is possible to infer whether the following phrase is related to ESG-related issues. In this study, we set the minimum number of words the model must read to 60. The CPT model could be used to calculate the probability that the phrase of interest is on an ERTP synonyms list. Notably, the probability of belonging is not intended to predict the next phrase because we do not intend to develop generative AI models like ChatGPT (Radford et al., 2018). The probability represents the propensity of the 60 input words to correspond to a particular ERTP. The central component of the CPT model is the self-attention block. Inputting 60 words into the CPT model for greenhouse gases, for instance, yields a probability of 0.85. We believe the pattern and a portion of the input words are strongly associated with greenhouse gases. The tendency of the entire annual report to a particular ERPT is the average value of all fragments extracted from the report.

The average probability predicted by the CPT model differs significantly from the frequency of phrases in annual reports. Counting phrase appearance directly is inflexible because counting requires precise word matching. Nevertheless, it is impossible to list all synonyms. Some phrases may be too ambiguous to be considered synonyms of a particular ERPT. The likelihood of a relationship could be a viable solution. According to the well-trained model, the output probability should be close to 1 if the phrase next to the input fragment is on the ERPT's list of synonyms, and vice versa.

CPT

The CPT models estimate whether a 60-word input fragment is associated with a particular ERPT. This is a typical task involving binary classification. The CPT model has eight layers: input layer, BERT embedding layer, transformer block layer, pooling layer, two dropout layers, and two dense layers (**Fig. 1**). The input layer creates a tensor, representing a 60-integer sequence. The embedding layer's primary function is to reduce tensor dimensions (Mikolov, Sutskever, Chen, Corrado, & Dean, 2013). The dictionary of the BERT tokenizer contains 30522 words. After one-hot encoding, each fragment becomes a 60×30,522 matrix, as each token is rewritten as a one-hot vector with 30,522 elements. The embedding layer reprojects the vectors into a 768-dimensional semantic space, significantly reducing the tensor size (Devlin et al., 2018). The transformer block layer, which is always at the center of the transformer model, employs a self-attention mechanism including an eight-head self-attention component (Vaswani et al., 2017) and a position-wise feedforward network. This layer has 18,946,592 parameters and assists the CPT model in softly focusing on potential GHG-related critical words and patterns (Vaswani et al., 2017). Averaging decreases dimensionality in the global average pooling layer. To avoid overfitting, two dropout layers are combined, and the dropout ratio is set to 0.1. To generate a number, two dense layers connect the network, each with 49,216 and 130 trainable parameters. A CPT model contains 128,478,178 parameters, including 18,995,938 trainable parameters and 109,482,240 non-trainable parameters derived from the BERT model. Moreover, we train the CPT models for 50 epochs with an early stopping patience of 20 epochs. Specifically, the model is trained for up to 50 epochs. If, after a certain period, the model's accuracy does not improve within the next 20 epochs, the training process is terminated before reaching the 50-epoch limit. We train and apply the network on four virtual machines on the Google Cloud Platform, each with an A100 40GB GPU, using TensorFlow 2.12.0 and Python 3.9.16. The GPU driver version is 525.105.17, and the CUDA version is 12.0. We use the Adam optimizer with a 0.0001 learning rate, categorical cross-entropy loss, and a 32-data batch size. Models are trained for 20 epochs and early stop with 5-epoch patience.

To create the training dataset, we downloaded and pre-processed a large number of annual reports from listed companies, with the requirement that the length of the annual report after tokenization not be less than 5,000 tokens. It should be noted that this annual report dataset contains investor reports; however, during training and validation, the training and validation datasets are strictly separated. Thus, the validation metrics are reliable. The processed dataset contains approximately 380 million tokens in total. The dataset could be divided into a similar number of 60-token fragments. All fragments are labeled based on the ERTPs found through ESS searches. For a specific ESG-related topic, if the two-word phrase following a fragment is in the corresponding ERTP list, the fragment is labeled as true; otherwise, it is labeled false. However, true-label data is relatively scarce. We randomly sample some data from all false data to keep the dataset size around 380,000. Following random selection, the dataset will be divided into two parts: the training dataset and the validation dataset. Their size ratio is roughly 9:1. We have 13 ESG-related topics, so we built 13 datasets to correspond with them.

We train 13 CPT models. Each CPT model is accountable for a specific ESG-related topic. In the experimental phase, we employ two training strategies for the CPT model. The current version is the first; training separated 13 CPT models. The alternative strategy is to train a unified model with 14 output categories, including a "None" category unrelated to any ERTPs. SoftMax is the final activation function of the unified model. Therefore, the probabilities of each ERTP could be estimated in a single prediction. The most significant benefit of training a unified model is that it can reduce training and report-analysis time. However, our dataset is unbalanced, making it difficult to determine the appropriate weights for each record. Additionally, some phrases may be associated with multiple ERTPs. Since the sum of SoftMax's outputs is always 1, the function's outputs would underestimate the relationship. Due to these factors, we opt for a separate-model plan.

Statistical indicators

Because all CPT models are for binary classification, the statistical indicators, including accuracy, precision, recall, and F1-score, are selected based on the confusion matrix. All statistical indicators are based on a 9:1 cross-validation ratio. In the 9:1 cross-validation, 90% of the records in the total dataset are used to train the CPT model, while the remaining 10% are used to test the model's accuracy. A confusion matrix summarizes the disparity between predicted and actual labels in the cross-validation dataset. **Table 1** is an example of a matrix of confusion. Actual labels are the output labels of the records: for a particular ERTP, the output label would be 1 if the raw output phrase is on the ERTP list, and 0 otherwise. Predicted labels are the labels of the CPT model prediction: if the predicted probability is greater than 0.5, the prediction is positive, and if it is less than 0.5, the prediction is negative. In the confusion matrix, a True Positive (TP) instance is one in which the actual label matches the predicted positive label; a False Positive (FP) instance is one in which the actual label does not match the predicted positive label; a True Negative (TN) instance is one in which the actual label matches the predicted negative label; and a False Negative (FN) instance is one in which the actual label does not match the predicted negative label. In the confusion matrix, the counts for TP, FP, TN, and FN instances are TP, FP, TN, and FN, respectively.

Accuracy is the percentage of instances correctly predicted out of the total number of instances. It provides an overall performance measurement of the CPT model, which is computed as follows:

$$Accuracy_j = \frac{TP_j + TN_j}{TP_j + FP_j + FN_j + TN_j} \tag{1}$$

where $Accuracy_j$ represents the accuracy of the jth CPT model for the jth ERTP, TP_j is the count of TP instances in the jth CPT model cross-validation, TN_j is the count of TN instances in the jth CPT model cross-validation, FP_j is the count of FP instances in the jth CPT model cross-validation, and FN_j is the count of FN instances in the jth CPT model cross-validation.

Precision is the proportion of correctly predicted positive instances relative to the total number of positive instances predicted. It emphasizes the precision of the CPT model's positive predictions, which are estimated as follows:

$$Precision_j = \frac{TP_j}{TP_j + FP_j} \tag{2}$$

where $Precision_j$ represents the precision of the jth CPT model for the jth ERTP. Recall measures the ratio of correctly predicted positive instances among the total actual positive instances, which is calculated as follows:

$$Recall_j = \frac{TP_j}{TP_j + FN_j} \tag{3}$$

where $Recall_j$ represents the recall of the jth CPT model for the jth ERTP. F1-score is the harmonic mean of precision and recall, providing a balanced measure of the model's performance. It considers both precision and recall and is especially effective when the dataset is imbalanced. F1-score is calculated as:

$$F1score_j = \frac{2 \times Precision_j \times Recall_j}{(Precision_j + Recall_j)} \tag{4}$$

where $F1score_j$ represents the F1-score of the jth CPT model for the jth ERTP.

ESG tendency analyses

After obtaining accurate CPT models, we analyze the available annual reports of investors. The annual report should be streamlined, tokenized, and broken down into discrete 60-word fragments. Each CPT model scans the datasets, which contain all the information required for an annual report. The CPT models generate 13 probability arrays for a single annual report. Each array is computed for a specific ERTP. We average the array of probabilities for a single investor during a specific year. Each word's relatedness to a specific topic in its context could be estimated by assuming that the probability of relatedness for each word in a fragment is the same as the probability of the fragment itself. Importantly, because the fragment moves one word forward at a time in the annual report and the fragment length is 60, the CPT models scan almost every word 60 times. As a result, each word has a probability of 60. We could calculate the relatedness probability of the word by averaging the 60 probabilities. The tendency of the annual report could be calculated using the relatedness probabilities of each word. In this article, the average probabilities are considered the ESG tendencies for each ERTP, ranging from 0 to 1, and are expressed as a percentage. The following equations use mathematical logic to estimate the tendency toward one topic in a single annual report:

$$Relatedness_{ij} = CPT_j(fragment_i) \tag{5}$$

where $fragment_i$ represents the i th fragment, $Relatedness_{ij}$ represents the relatedness level between the ith fragment and the jth topic, and CPT_j represents the CPT model corresponding to the jth topic. The tendency of an annual report is computed as follows:

$$Tendency_{jk} = \sum_{i=0}^{n_k} \frac{Relatedness_{ijk}}{n_k} \tag{6}$$

where $Tendency_{jk}$ represents the kth annual report's tendency toward the jth topic, n_k is the number of fragments that could be encoded from the kth annual report, and $Relatedness_{ijk}$ is the $Relatedness_{ij}$ using the kth annual report.

Importantly, the CPT method differs significantly from the word-counting method. The word-counting method counts the words on a predetermined dictionary and then calculates the word frequency in the entire annual report. First, the word-counting method necessitates that the ERTP dictionary should be exact. The output of the word-counting method is either 1 or 0; there is no intermediate state. When some phrases are ambiguous, the word-counting method yields irrational results. The CPT technique is more adaptable. The CPT method estimates the likelihood of a phrase having a close relationship to an ERTP. Second, the output probability from the CPT method represents the tendency of the 60-word input fragment rather than a single phrase.

We also look at the annual trends in tendency changes, the differences between investors from different sub-styles, and the disparity between investors from the six major countries with the highest GDPs, namely China, France, Germany, Japan, the United Kingdom, and the US. This study has a global scope and includes countries like Switzerland, South Korea, Liechtenstein, and Canada. Due to space constraints, we chose the largest country as a representative sample. To highlight the differences between different sub-styles and counties, we standardize all tendency values as follows:

$$StanTendency_{jk} = \frac{Tendency_{jk} - MeanTendency_j}{StdTendency_j} \quad (7)$$

where $StanTendency_{jk}$ represents the standardized value of the kth annual report's tendency toward the jth topic, $MeanTendency_j$ is the average value of all reports' tendency toward the jth topic, and $StdTendency_j$ is the standard deviation of all reports' tendency toward the jth topic.

Results

Statistical indicators of CPT models

The statistical indicators of the CPT models for each ERTP are displayed in **Table 2**. The CPT model for the ERPT, "governance risk" has the lowest accuracy, 94.41%, and the highest F1-score, 92.28%. The ERPT "governance risk" has the most instances labeled as 1 in the training dataset. According to relatively similar accuracy, precision, recall, and F1-score values, the "governance risk" CPT model for the ERPT is the most reliable among the 13 trained CPT models. The CPT model for the ERPT, "domestic reflux rate," has the highest degree of accuracy, 99.72%, and F1-score is also good, 91.99%. The number of synonym phrases in the ERPT synonym list affects the performance of the model. The ERPT synonym lists contain 24 and 9 synonyms for "governance risk" and "domestic reflux rate," respectively. On the contrary, the CPT model for the ERPT question "domestic job creation" has the poorest performance. The accuracy, precision, recall, and F1-score are 98.10%, 42.44%, 70.78%, and 53.06%, respectively. The number of records that contain positive label in its cross-validation dataset is 959, or 2.54%, but its ERPT synonym list contains 60 terms. In other words, the entire dataset contains few phrases for each synonym. It is difficult for the CPT model to precisely "remember" phrases that appear less than five times in the total dataset, and the model's profile learned from the dataset with relatively scattered ERTPs is also vague. During cross-validation, some records that appear less frequently in the training dataset are difficult to recognize, and many phrases that are not on the ERTP list of "domestic job creation" are incorrectly predicted as positive instances. As a result, the model's precision decreases. The CPT model for ERPT, "work environment," is similar in that its precision is lower because the ERTP list for "work environment" is also relatively large and sparse. On the basis of numerous experiments, we balance training data, model construction, and training strategies. In conclusion, the current model is the best option under current conditions.

Average ESG tendency of all annual reports

Table 3 shows the average ESG tendency across all annual reports in our database. Out of 13 ESG-related topics, production cost, and governance risk have the highest tendency, accounting for 10.95% and 10.02%, respectively. These annual reports pay least attention to the domestic reflux rate, which is 0.68%. Investors place a higher value on water consumption, accounting for 2.53%, while safety and health receive the highest priority, accounting for 4.72%. The average trends for environmental, social, and economic factors are 2.19%, 2.74%, and 8.51%, respectively. For these investors, the economic aspect is of the utmost importance, whereas the environmental, and social aspects receive equal attention.

Temporal variation in ESG tendency

Table 4 summarizes the ESG tendencies by year. There are approximately 200 annual reports per year, a number that remains stable over time. According to the mean values of ESG tendencies in all available annual reports, among 13 ERTPs, the majority of investors' language in the annual reports relates to "production cost" and "governance risk" annually. These two subjects consistently stand out as pivotal components of investors' annual reports. Conversely, the "domestic reflux rate" has consistently garnered less emphasis. **Fig. 2** depicts the changes in ESG trends for each ERTP throughout the year. Investors gradually increase the weights of ESG tendencies of ERTPs, including "greenhouse gas" (**Fig. 2b**), "work environment" (**Fig. 2e**), "safety and health" (**Fig. 2g**), "community" (**Fig. 2f**), "human rights" (**Fig. 2h**), "domestic job creation" (**Fig. 2i**), "domestic reflux rate" (**Fig. 2j**) and "governance risk" (**Fig. 2k**). According to **Figs 2a**, **2c**, **2d**, **2l**, and **2m**, the ESG trend of other ERTPs, including "air pollution," "water consumption," "mining consumption," "production cost," and "economic ripple effect," is decreasing over time.

Differences in ESG tendency among sub-styles

We divide the investors into eight sub-styles based on the GICS (MSCI, 2023b), which are "bank and trust," "corporation," "hedge fund," "holding company," "insurance company," "investment advisor," "investment advisor/hedge fund," and "research firm." Specifically, the GICS divides all investors into three styles--investment managers, strategic entities, and brokerage firms--and then into eight sub-styles. We prefer the sub-style category because it is more intuitive and clear. **Table 5** shows the tendency for each sub-ESG style. To make the distinction between the sub-style investors clear, **Fig. 3** shows the average standardized values for each topic, organized by sub-styles.

Table 5 shows that almost all investors place great importance on governance risk and production cost. Compared to other sub-styles, "bank and trust" investors most frequently mention the content related to "community," "economic ripple effect," and "domestic reflux rate," which are 0.27, 0.31, and 0.12 standard deviations above the averaged values, respectively (**Fig. 3a**). The corporation investors, focusing on entities, express the most attention to "air pollution," "greenhouse gas," "water consumption," "mining consumption," "work environment," and "human rights," which are 0.30, 0.30, 0.32, 0.24, 0.22, and 0.11 standard deviations higher than the average values, respectively (**Fig. 3b**). However, "corporation" investors notice risk governance the least. The ESG tendencies of "hedge fund," "investment advisor," and "investment advisor/hedge fund" investors are similar, except for "economic ripple effect" and "governance risk," as their tendencies toward other topics are lower than the average levels (**Figs. 3c**, **3f,** and **3g**). Interestingly, "holding company" investors are the most interested in safety and health, roughly 0.91 standard deviations higher than the average value, followed by "insurance company" investors approximately 0.34 standard deviations higher than the average level (**Figs. 3d** and **3e**). Except for "research firm" investors (6.27%), the production cost tendencies of all other investors are roughly 10% as shown in **Table 5**. "Insurance company" and "corporation" investors focus the most on production cost, around 0.16 and 0.12 standard deviations higher than the average values, respectively (**Figs. 3e** and **3b**). Compared with other investors, "research firm" investors focus on governance risk and domestic job creation the most, compared with other investors, which are 1.33 and 0.35 standard deviations over the average levels, respectively (**Fig. 3h**).

ESG tendency temporal variation in different classes

The sub-style is further divided into two classes: the financial class and the nonfinancial class. The nonfinancial class consists primarily of corporation and holding company investors, whereas the financial class includes all other investors. The nonfinancial class pays more attention to environmental and human-related issues, such as air pollution, greenhouse gases, water consumption, mining consumption, work environments, safety, and health, and human rights, than the financial class.

Fig. 4 displays the ESG tendency temporal variation in different classes from 2010 to 2021. The trends of temporal variation of greenhouse gas (**Fig**. **4b**), work environments (**Fig. 4e**), community (**Fig. 4f**), safety, and health (**Fig. 4g**), human rights (**Fig. 4h**), domestic job creation (**Fig. 4i**), and economic ripple effect (**Fig. 4j**) in the financial and nonfinancial classes are roughly the same, although the degrees vary. Nonfinancial investors are more concerned with environmental issues, such as air pollution, water consumption, and mining consumption (**Fig. 4a**, **4c**, and **4d**). However, it appears that the attention paid to these issues is gradually diminishing. The investors in the financial class continue to hold a relatively stable and subdued view of these topics. In recent years, financial investors have increased their emphasis on governance risk, while nonfinancial investors have become less concerned (**Fig. 4k**). Both financial and nonfinancial investors gradually mention a decline in production costs, but in 2021, only nonfinancial investors increase to emphasize it (**Fig. 4l**). Beginning in 2018, nonfinancial investors are more aware of the domestic reflux rate, whereas financial investors are less aware (**Fig. 4j**).

Differences in ESG tendency among major countries

Fig. 5 shows the standardized ESG orientation of investors in the six countries with the highest GDPs, excluding India. Importantly, the top 350 investors come from all over the world and not just these six countries. There are no Indian investors among the top 350 investors with the highest investment ownership. Investors in China are most concerned with production costs. The least attention is paid to air pollution, greenhouse gas countries, water consumption, mining consumption, and economic ripple effects compared to foreign investors (**Fig. 5a**). Investors in France are more concerned with human rights than investors in other countries, while they rarely focus production cost and domestic reflux rate (**Fig. 5b**). Investors in Germany pay the most attention to safety and health and economic ripple effect, while they pay the least attention to community (**Fig. 5c**). The Japanese investors emphasize community, air pollution, greenhouse gas, water consumption, work environment, domestic job creation, and domestic reflux rate, but governance risk is mentioned the least (**Fig. 5d**). Investors in the United Kingdom are more concerned with governance risk than they are with safety and health (**Fig. 5e**). United States investors emphasize mining consumption the most. In contrast, they make the fewest references to work environment, human rights, and domestic job creation (**Fig. 5f**).

In terms of ESG tendency patterns, the United States and China are comparable (**Fig. 5a** and **5f**). They pay relatively little attention to environmental issues such as water consumption, greenhouse gas emissions, and air pollution. Although these topics' ESG tendencies are the lowest among Chinese investors, their ESG tendencies among American investors closely follow those of Chinese investors. The three European countries' patterns are similar (**Fig. 5b**, **5c**, and **5e**). Japan shows the most prominent trend in both environmental and social issues (**Fig. 5d**).

ESG tendency temporal variation in major countries

From 2010 to 2021, **Fig. 6** depicts the ESG trend of investors in the six largest countries. Clearly, we add linear fitting lines to show the changing trend. Japan's investors are keenly interested in the majority of ESG-related topics. However, attention to air pollution, water consumption, and mining consumption decreased over time (**Figs. 6a**, **6c**, and **6d**), despite the current value being the highest among the six counties. Investors in the United Kingdom are discussing all ESG-related topics with the exception of production cost more frequently. The situation is comparable to that of German and French investors. Investors in the United States are gradually emphasizing mining consumption and governance risk, while remaining relatively stable on other issues. Investors in China are gradually increasing their focus on safety and health, human rights, governance risk, and greenhouse gas emissions.

Discussion

This study is a significant step forward in research because it pioneers the use of NLP and machine learning to uncover trends on the 13 ESG-related topics from unstructured data, specifically investor annual reports. We meticulously analyzed 2,533 annual reports from 2010 to 2021, sourced from the top 350 investors with the highest ESG ownership. We processed over 102 million valid words using our computational models, which would have been extremely difficult for human analysts to do on their own. Our findings show that investor attention has shifted away from purely economic metrics toward environmental and social considerations since 2010, though the economic aspect remains the most important. We observed that investors' preferences tend to align with their respective investor types' preferences. Furthermore, our findings show a consistent evolution of ESG tendencies in financial and nonfinancial investors over time. We also discovered differences in investor tendencies by country, as well as variations in trends across countries. To facilitate this research, we present CPT, a novel textual analysis methodology that effectively and objectively analyzes the language of the annual report on ESG topics. Overall, this study provides valuable insights into investors' changing priorities and emphasizes the importance of using advanced computational methods to extract meaningful information from large amounts of textual data.

In terms of overall tendency composition, investors continue to prioritize economic aspects; however, they are gradually shifting their focus to social and environmental aspects globally, consistent with single-country studies (Park & Jang, 2021; Sood et al., 2023; Tang et al., 2024). The reasons for this change and pattern are numerous. First, governments and organizations develop and mandatorily implement more environmental and social regulations (Baumüller & Sopp, 2022; Gustafsson et al., 2023; Linsley, Abdelbadie, & Abdelbadie, 2023). For example, the United States Greenhouse Gas Reporting Program, which began in 2010, required firms to disclose their GHG emissions and affected 6200 facilities, whose emissions roughly accounted for half of total US emissions in that year (Tomar, 2023). Furthermore, as a result of the criminalization of human rights violations under the French Duty of Vigilance law, companies, particularly transnational companies, are paying more attention to human rights as a result of the criminalization of human rights violations (Gustafsson et al., 2023; Schilling-Vacaflor, 2021). The Non-Financial Reporting Directive has institutionalized and standardized ESG-related information disclosure (Baumüller & Sopp, 2022) in the European Union. Second, the investors generally avoid environmental and social risks (Bolton & Kacperczyk, 2021; Cornell, 2021). In other words, investors focus more on these risk factors. In relation to our results, the significance of the GHG emissions gradually increases; for example, Han, Lee, and Wang (2023) show that foreign investors avoid investing in Korean firms with high GHG emissions. Third, in terms of the importance, the economic aspect always receives the most attention, which makes sense as our materials are the annual reports. Annual reports are financial reports that primarily disclose information about business conditions and strategies (Ramzan et al., 2021; Stanton & Stanton, 2002). To pursue greater profitability and efficacy, companies inevitably emphasize production cost, governance risk, and economic ripple effect. Our results are consistent with this fundamental premise.

Several factors influence the variation in ESG tendencies among investors' sub-style businesses. On the one hand, nonfinancial investors, specifically corporations and holding company, are closer to the production, indicating that there will be more laws and regulations that will directly impact them. For example, China's mandatory disclosure regulation requires the heavily polluted firms to disclose their environmental impacts (Z. Zhang, Su, Wang, & Zhang, 2022), despite the fact that all of these firms are nonfinancial sectors. Nonfinancial investors are more likely to disclose information about their environmental and social impacts to reduce their systematic risk and cost of equity as governments focus more on the commodity and supply chain (Cuomo, Gaia, Girardone, & Piserà, 2022). Our results indicate that "corporation" investors are significantly prone to disclosure more likely to disclose environmental and social information, consistent with the context and logic of previous literature (Cuomo et al., 2022; Z. Zhang et al., 2022). Financial investors, on the other hand, tend to reduce governance risk because of their critical role in maintaining the integrity, stability, and compliance of financial institutions (Aevoae, Andrieş, Ongena, & Sprincean, 2023; Di Tommaso & Thornton, 2020), and they have fewer less direct impacts on environmental and social aspects and indirectly mediate them through green financing (X. Zhang, Wang, Zhong, Yang, & Siddik, 2022). Consequently, there is a substantial difference in sub-styles between financial and nonfinancial investors. Furthermore, recently, both financial and nonfinancial investors have been gradually increasing their emphasis on key topics such as GHG, human rights, and the work environment (Bolton & Kacperczyk, 2021; Han et al., 2023). Nonfinancial investors are primarily responsible for variations in overall tendencies toward air pollution, water consumption, and mining consumption because they are most impacted and supervised the most in these areas by governments and society. Regarding the time-series changes in the trend toward ESG in various countries, only on the topic of GHG do all of the changes point in the same direction.

Several factors contribute to the variation at the country level. It is important to note that our scope is global, encompassing more than 20 countries and regions, and we only discuss the six countries with the highest GDP. Voluntary ESG reporting is a major reason why Japanese investors' ESG tendencies differ significantly from those of other countries (Nakajima & Inaba, 2022). Around 2015, Japanese investors and firms began replacing traditional annual reports with integrated reports (Oshika & Saka, 2017). According to the reports in our dataset, investors from other countries do not use integrated reports as widely and extensively, and instead conduct nonfinancial disclosure through separate nonfinancial reports. Although the patterns of the United States and China are comparable, their characteristics are distinct. The majority of financial investors in the United States prefer to pay greater attention to governance risk and production cost. As the largest developing country, China's investors are expanding, and seeking a larger market share and better financial performance; therefore, economic factors are the most important consideration. Japan lacks both natural and human resources, and corporations account for the majority of its investors. Because they are closer to the production line, they naturally prioritize environmental and social concerns. Prior to Brexit, German, French, and UK investors were primarily governed by various EU agreements and mandatory disclosure regulations, such as the Non-Financial Reporting Directive (Baumüller & Sopp, 2022; Caputo, Pizzi, Ligorio, & Leopizzi, 2021). Moreover, these three countries have similar levels of economic development. Therefore, they represent a new paradigm. Another factor that may contribute to country-level variation is the disparity in accounting standards between nations. Although these standards may influence textual content, there are currently no methods to address this disparity. However, many countries worldwide are in the process of adopting new standards proposed by the Sustainability Accounting Standards Board, which will go into effect between 2023 and 2025. This could be a critical step toward resolving the issue.

Multiple potential stakeholders, including governments, companies, investors, and consumers, can derive valuable insights and implications from our research. Governments can leverage our findings to comprehend the average level of investor tendencies on various ESG topics across several major countries. Additionally, by understanding the average tendencies of different investor sub-styles, governments can formulate more targeted policies to guide these diverse investor groups (Liu, Cifuentes-Faura, Zhao, & Wang, 2024). For companies, a deeper understanding of investor preferences can enhance their prospects of attracting investment. This insight may encourage companies to adopt ESG-related strategies more proactively and to disclose ESG-related information more transparently (Chen & Xie, 2022; Van Duuren, Plantinga, & Scholtens, 2016). Our study highlights the ESG tendencies of the largest investors, indicating that other investors may need to adjust their strategies to remain competitive (Assaf, Monne, Harriet, & Meunier, 2024). Investors often play a pivotal role in market activities, and their emphasis on ESG considerations will eventually influence consumers through the products and services offered by the invested companies. Consumers can thus gain an understanding of market trends by analyzing investor behavior and take informed actions to safeguard their interests.

Our CPT models substantially improve textual analyses of annual reports. Textual analyses are essential in ESG evaluations (Baier et al., 2020; Giles & Murphy, 2016; Lokuwaduge & Heenetigala, 2017; Loughran & McDonald, 2016). The methods in previous studies can be roughly divided into two categories: the first way is to define a dictionary of ESG and then count the occurrences of each element of the materials, e.g., Baier et al. (2020), Jiang, Gu, and Dai (2023) and Loughran, McDonald, and Otteson (2023); the second approach is to manually classify the content through human reviews of the reports and based on the researchers' knowledge, e.g., D'Augusta, Grossetti, and Imperatore (2023), Giles and Murphy (2016), and Lokuwaduge and Heenetigala (2017). The fundamental premise of the appearance-counting method is that documents will mention a subject more frequently if their authors have a vested interest in it. The first approach requires fewer resources and is more objective, but it ignores semantic context. Furthermore, deciding whether a word or a phrase should be included in the dictionary is difficult because some phrases are only partially related to a topic. The second approach focuses on human resources as the critical bottleneck, but the analyses of each sentence are more thorough. Furthermore, this method has a significant degree of bias and subjectivity. Previous studies, such as Giles and Murphy (2016) and Tilling and Tilt (2010), have used a second-order analysis scale of less than

1000 reports. The CPT models consider the benefits of both methods at the cost of computing resources. Our method decreases reliance on the predefined word lists, specifically ERTPs in this study, and broadens the relatedness output scale from a binary 0 or 1 to a continuous range of 0 to 1. For instance, “sulfur dioxide” appears infrequently mentioned in our training dataset and is associated with air pollution, but it is not included in the ERTP. Thus, on the basis of preceding 60-word fragment, it should not be considered relevant. Nonetheless, our trained CPT model predicts a 0.8677 output in an experimental instance, which contradicts human judgment, because the prediction is based on the semantic context of the input 60-word segment. This shows how the CPT models correct human inaccuracies. Unlike previous studies, our dictionary’s purpose is to assist models in determining which fragments are relevant to the topics of interest. Due to probability-based correction, we can use reasonably sized dictionaries while ensuring that focus is not missed, which is an improvement over previous studies (Baier et al., 2020; Bodnaruk et al., 2015). In other words, we only need to add common and frequently used terms to the dictionary of CPT models, which significantly reduces the difficulty of creating it. Because we use CPT models to compute the relatedness probabilities of fragments, the results of this method differ from appearance-counting statistics performed directly on the dictionary. Although they primarily digest sentences from reports, the 60-word fragments may retain the semantic contexts as seen in previous studies. We do not analyze sentences directly because it is difficult to accurately extract sentences from PDF files. Furthermore, models based on the transformer block require fixed input lengths (Vaswani et al., 2017). Although we could use the padding and punctuation technology to manipulate the sentences, a 60-word fragment may be more efficient and effective. This allowed our models to objectively analyze large-scale datasets with high accuracy and low cost.

FinBERT (Huang et al., 2023), KoBERT (Bang et al., 2023), and ESGBERT (Mehra et al., 2022) are examples of AI models developed for report sentiment analysis and ESG content classification. The tasks of the CPT models are not the same as those in previous studies, that is, the CPT models are to estimate the relatedness between each fragment in a report and a specific topic related to ESG; FinBERT is to analyze the sentiment of the report sentences (Huang et al., 2023); KoBERT is to judge whether Korean content is related to ESG (Bang et al., 2023); and ESGBERT is to probe the change in environmental risk score change or not and the direction of change (Huang et al., 2023; Mehra et al., 2022). Furthermore, in terms of accuracy, the CPT models outperform FinBERT’s 88.2% and ESGBERT’s 67.09% for the change classification task and 79.30% for the change direction classification. However, zero-shot learning is gradually gaining popularity. For example, a previous study showed that the fine-tuned GPT-2 model could achieve an average 94% accuracy in multilabel classification tasks (Puri & Catanzaro, 2019). Additionally, fine-tuning more advanced models, such as ChatGPT 4.0, Meta Llama 2, and Google PaLM, among many others, may improve performance even further. However, for specific problems and topics, CPT models should perform better than large models because they can handle such tasks with fewer resources, specifically fewer GPUs and lower electricity consumption.

This study has several limitations. First, due to our current computing limitations, we only cover 350 investor annual reports from 2010 to 2021, despite the fact that the size of the investigated data totals 102 million. A larger data set could improve the CPT model’s precision and adaptability and investigate for a more thorough examination of country-level differences. In the words, more investors should be included in the analysis. According to annual reports, there is a certain amount of delay. The majority of annual reports become available the following year. Whether other official resources could be utilized is still up for debate. Third, despite the impressive performance of the current CPT models, additional fine-tuning, and training are necessary to improve their performance. Moreover, a certain level of sampling bias problem exists in our study. Since LLM is expensive to train and use, we could not analyze thousands of investors. Therefore, we selected 350 investors based on previous research. These 350 investors have the largest ESG ownership, and they also have relatively large capitalization. In this way, their tendency variation would cause a ripple effect in leading responsible investments. All of them are large and relatively more transparent and have thousands of investments either green or brown. However, our research has relatively neglected small investors and investors who mainly invest in non-listed companies. In future research, more unstructured data will be collected and analyzed to examine the temporal variation of ESG studies in greater detail. More investors could be considered to reduce sampling bias. Furthermore, over time, language evolves, and popular terms undergo changes. It is crucial to delve deeper into the fluctuations in the usage of specific words to depict shifts in institutional language and societal trends. Improvements to the CPT models would be considered with a larger data set and better architecture. Unknown is whether the ESG trend impacts companies' decision-making and allows them to act. Consequently, it is necessary to investigate the connection between their tendencies and actions. Finally, while multicollinearity can occur in deep learning models, its impact is managed through the inherent capabilities of neural networks to learn complex patterns, the use of regularization techniques, and dimensionality reduction methods. However, it remains important to be aware of multicollinearity.

Conclusion

Based on a textual analysis of more than 3,000 annual reports released between 2010 and 2021, the general tendencies of investors toward diverse ESG-related topics, their temporally evolving pattern, the difference between investors from various sub-style businesses, and the potential variation of the composition of investor inclinations across major countries are investigated in this study. This study shows that investors are most concerned with the economic aspects of ESG, specifically governance risk, and production cost. However, their emphasis on production costs diminishes over time. In addition, investors are increasing a greater importance on ESG factors. Environmental and social factors are mentioned more frequently by nonfinancial investors than by financial investors. Japan's investors place a greater emphasis on the environmental and social aspects of ESG compared to investors from China and the United States, whose annual reports focus less mention of these countries. Our results suggest that investors are increasingly recognizing the importance of sustainability considerations in their decision-making process. Investors in various businesses may face different ESG challenges and opportunities, necessitating targeted approaches to address them effectively. Companies may need to improve their ESG disclosures to align with changing investor preferences and increasing transparency and accountability as investors increasingly focus on environmental and social factors. In terms of rules and regulations, mandatory ESG-related disclosure regulations should require firms to disclose not only the amount of environmental and social impacts but also the implementation of the discloser, such as ways to reduce the negative impacts and future strategies. Recognizing the differences in ESG challenges and opportunities across industries, regulators can develop sector-specific reporting standards tailored to the distinct characteristics and needs of various business sectors. Given the global nature of capital markets and the differences in ESG reporting practices across countries, regulators may work with international organizations and standard-setting bodies to harmonize ESG reporting standards and frameworks.

Acknowledgments

This research was supported by the following funding agencies: JSPS KAKENHI (Grant No. JP20H00648 and No. JP21K17927), the Environment Research and Technology Development Fund of the Environmental Restoration and Conservation Agency of Japan (Grant No. JPMEERF20201001), and JST MiraiProgram (Grant No. JPMJMI22I4).

Tables

Table 1. Confusion Matrix

		Actual Label	
		1: Positive	0: Negative
Predicted Label	1: Positive	TP	FP
	0: Negative	FN	TN

Note: The label is for ERTP.

Table 2. Statistical Indicators of CPT Models

ERTP	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	TP	TN	FP	FN
Air Pollution	98.50	61.41	79.78	69.40	643	36620	404	163
Greenhouse Gas	97.61	89.37	87.29	88.32	3421	33504	407	498
Water Consumption	98.35	68.81	82.78	75.15	942	36265	427	196
Mining Consumption	98.93	78.04	87.06	82.31	942	36483	265	140
Community	98.37	77.14	79.30	78.21	1107	36106	328	289
Work Environment	97.60	59.66	66.26	62.79	766	36156	518	390
Safety and Health	96.31	81.77	73.10	77.19	2364	34069	527	870
Human Rights	98.14	78.04	78.71	78.37	1272	35856	358	344
Domestic Job Creation	98.10	42.44	70.78	53.06	407	36703	552	168
Domestic Reflux Rate	99.72	87.90	96.48	91.99	603	37122	83	22
Governance Risk	94.41	95.70	89.09	92.28	12632	23083	568	1547
Production Cost	97.85	84.46	87.76	86.08	2517	34499	463	351
Economic Ripple Effect	98.27	69.11	87.17	77.10	1101	36075	492	162

Table 3. Average ESG Tendency of All Annual Reports

Aspect	ESG-related Topic	ESG Tendency (%)	Aspect Average (%)
Environmental Aspect	Air Pollution	1.79	2.19
	Greenhouse Gas	2.20	
	Water Consumption	2.53	
	Mining Consumption	2.23	
Social Aspect	Work Environment	2.73	2.74
	Community	2.85	
	Safety and Health	4.72	
	Human Rights	2.00	
	Domestic Job Creation	3.44	
	Domestic Reflux Rate	0.68	
Economic Aspect	Governance Risk	10.02	8.51
	Production Cost	10.95	
	Economic Ripple Effect	4.55	

Table 4. ESG Tendency Summary by Year

Year	Air Pollution (%)	Greenhouse Gas (%)	Water Consumption (%)	Mining Consumption (%)	Community (%)	Work Environment (%)	Safety and Health (%)
2010	2.22	2.29	2.68	2.71	2.14	2.28	4.55
2011	2.02	2.11	2.97	2.66	2.79	2.48	4.52
2012	1.99	2.05	2.97	2.61	2.65	2.38	4.98
2013	1.45	1.83	2.33	2.35	2.28	2.27	4.12
2014	1.86	1.85	2.59	2.02	2.25	2.16	4.49
2015	1.58	1.89	2.15	2.55	2.99	2.67	4.54
2016	2.29	1.90	3.09	2.33	2.86	2.67	4.63
2017	1.71	1.86	2.31	1.84	3.10	2.96	4.65
2018	1.95	2.12	2.72	2.23	3.16	3.05	4.78
2019	1.70	2.23	2.55	2.32	3.17	2.93	4.75
2020	1.51	2.69	2.22	1.67	3.52	3.50	5.35
2021	1.28	3.49	1.94	1.52	3.28	3.41	5.27
Total	1.79	2.20	2.53	2.23	2.85	2.73	4.72

Year	Human Rights (%)	Domestic Job Creation (%)	Domestic Reflux Rate (%)	Governance Risk (%)	Production Cost (%)	Economic Ripple Effect (%)	Report Count (%)
2010	1.33	2.87	0.91	9.10	11.53	5.16	212
2011	1.46	3.48	0.15	9.24	11.55	5.49	198
2012	1.55	3.19	0.39	9.35	11.72	4.84	206
2013	1.41	2.96	0.60	9.79	12.00	4.41	217
2014	1.63	3.07	1.27	9.81	11.76	4.14	218
2015	2.05	3.62	0.40	10.09	11.23	4.57	211
2016	1.92	3.54	0.49	9.48	11.54	4.58	202
2017	2.03	3.69	0.65	10.37	10.57	4.61	219
2018	2.39	3.70	0.81	10.23	10.63	3.88	194
2019	2.23	3.59	0.56	10.56	10.18	4.09	222
2020	2.99	3.80	1.01	10.99	8.96	4.53	203
2021	2.96	3.83	0.89	11.11	9.77	4.30	231
Total	2.00	3.44	0.68	10.02	10.95	4.55	2533

Table 5: The Average Tendency of Each Sub-Style

	Bank and Trust	Corporation	Hedge Fund	Holding Company	Insurance Company	Investment Advisor	Investment Advisor/Hedge Fund	Research Firm
Air Pollution (%)	0.44	3.13	1.10	1.36	0.36	0.59	0.47	0.29
Greenhouse Gas (%)	1.23	3.10	0.77	2.24	0.75	1.39	1.53	1.50
Water Consumption (%)	0.64	4.23	1.30	2.38	0.39	1.10	0.95	0.50
Mining Consumption (%)	0.98	3.40	1.78	0.95	0.88	1.21	1.48	0.64
Work Environment (%)	1.97	3.29	1.39	2.66	1.94	2.36	2.18	1.97
Community	3.82	3.01	1.68	1.95	1.70	2.51	1.62	1.95
Safety and Health (%)	1.88	6.09	3.10	9.25	6.43	3.84	3.13	1.34
Human Rights (%)	1.40	2.26	0.99	2.12	1.44	1.90	1.96	1.65
Domestic Job Creation (%)	3.75	3.86	2.89	4.44	2.59	2.78	2.09	4.47
Domestic Reflux Rate (%)	1.21	0.83	0.06	0.36	0.06	0.20	0.59	1.09
Governance Risk (%)	16.08	6.39	11.47	7.30	10.61	12.34	13.33	19.42
Production Cost (%)	9.87	11.78	10.86	10.31	12.03	10.34	10.31	6.27
Economic Ripple Effect (%)	5.88	3.76	4.43	4.56	3.04	5.23	4.74	4.98

Figures

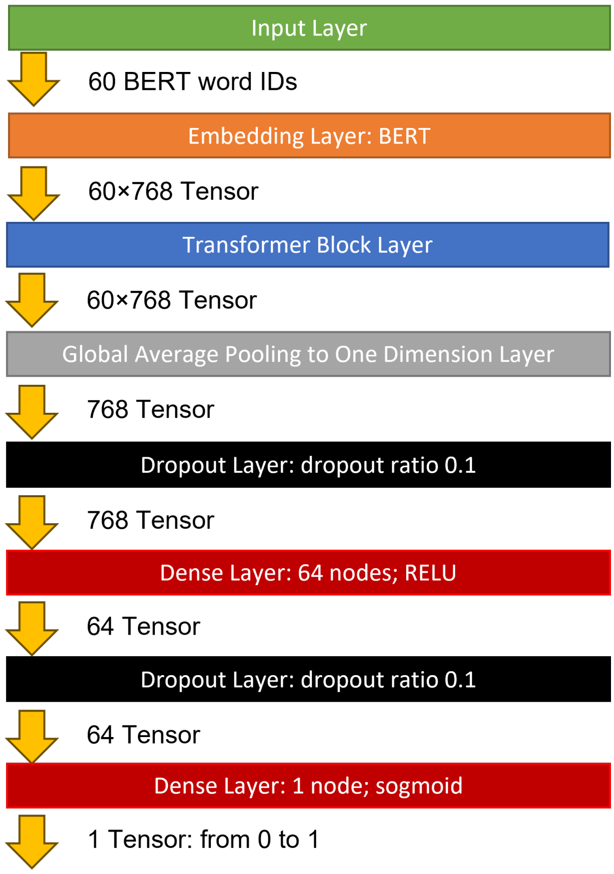

Fig. 1. Network Structure of CPT

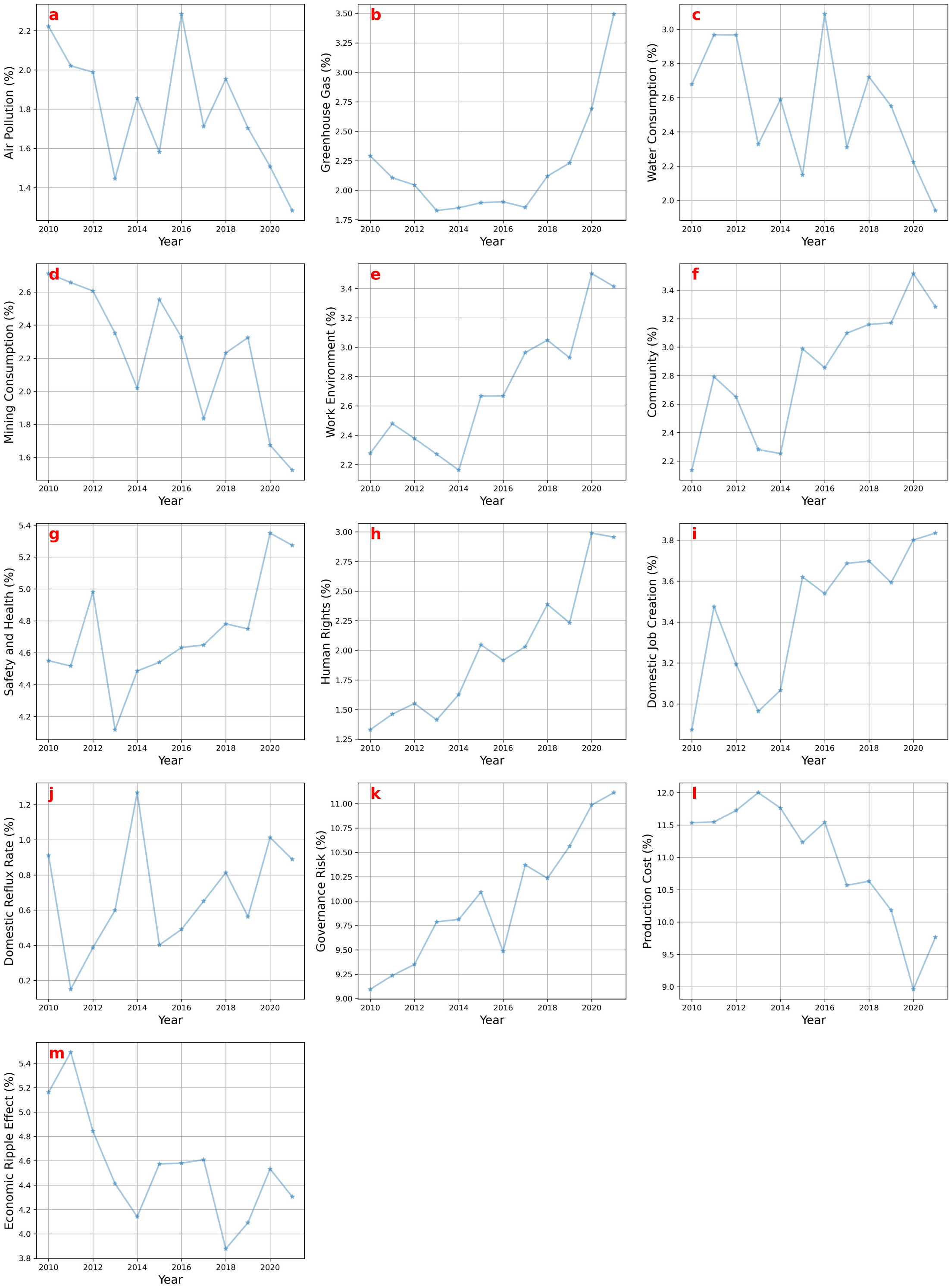

Fig. 2. The Variation in ESG Tendency across Years

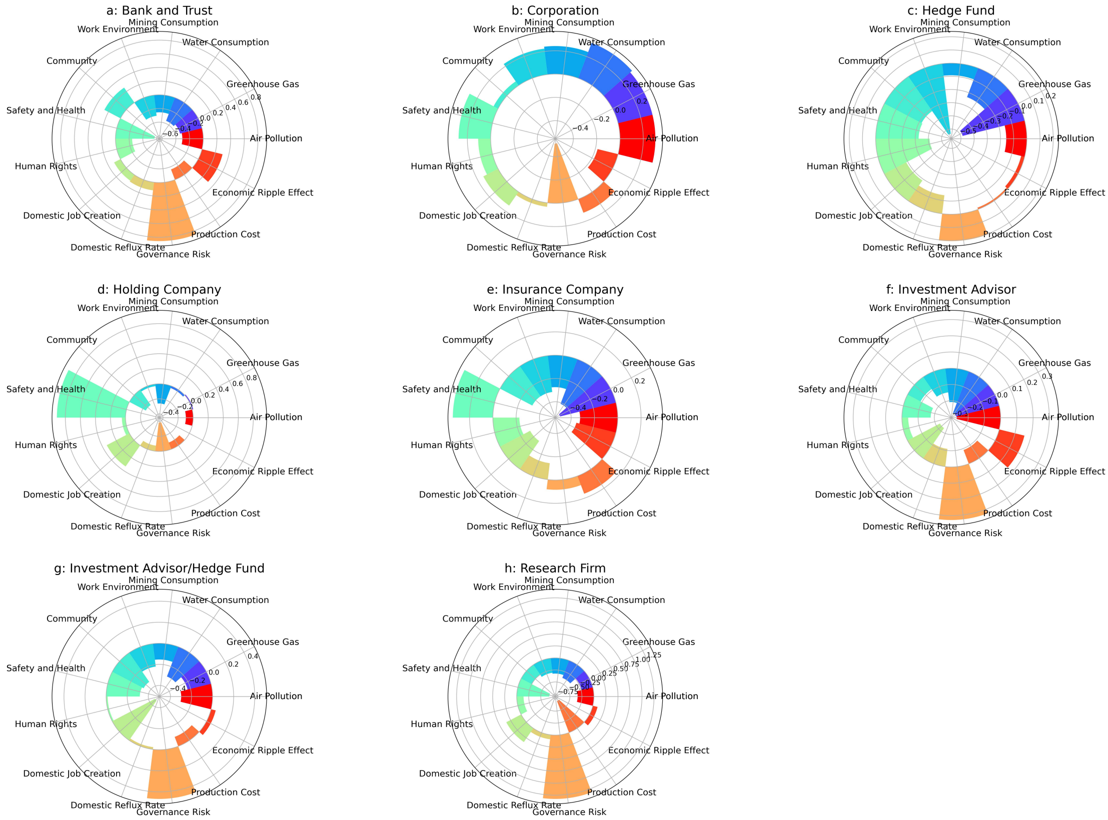

Fig. 3. The ESG Tendency of Each Sub-Style

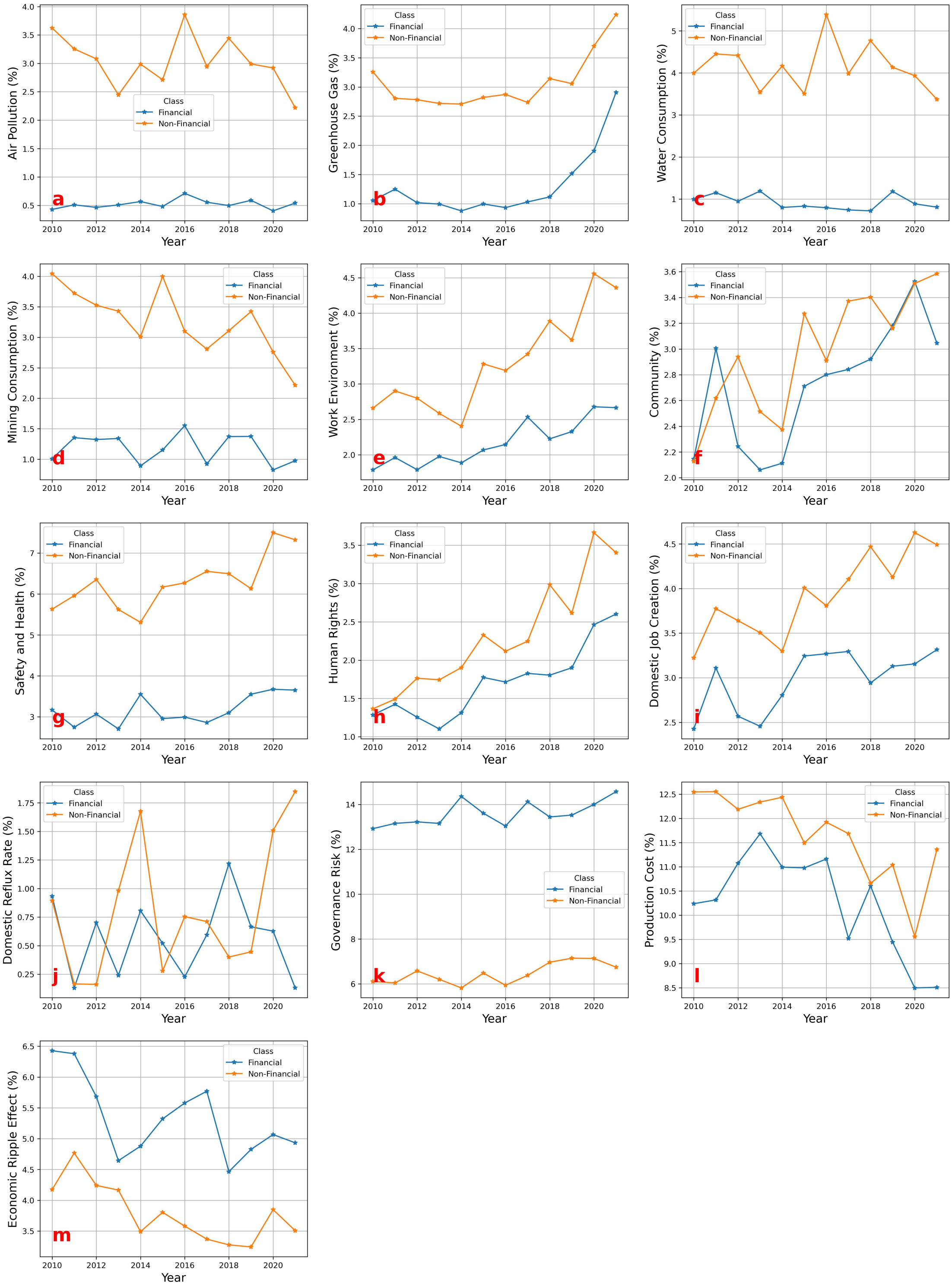

Fig. 4. The ESG Tendency Temporal Variation of Two Classes

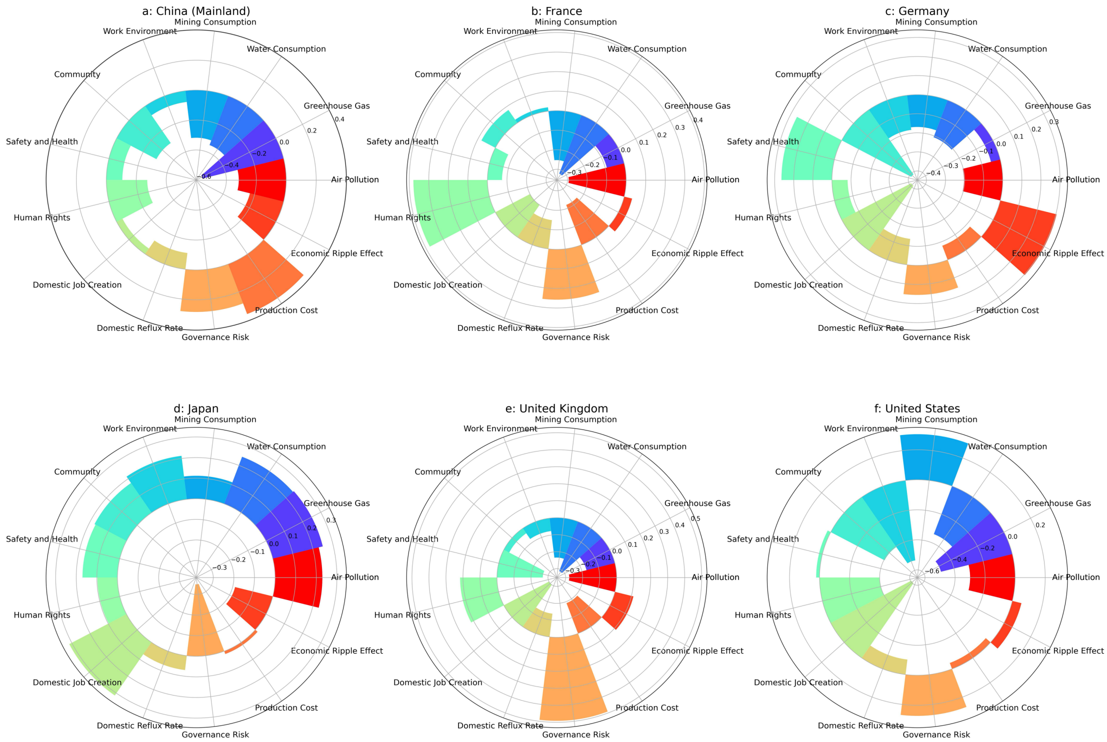

Fig. 5. The ESG Tendency of Major Countries

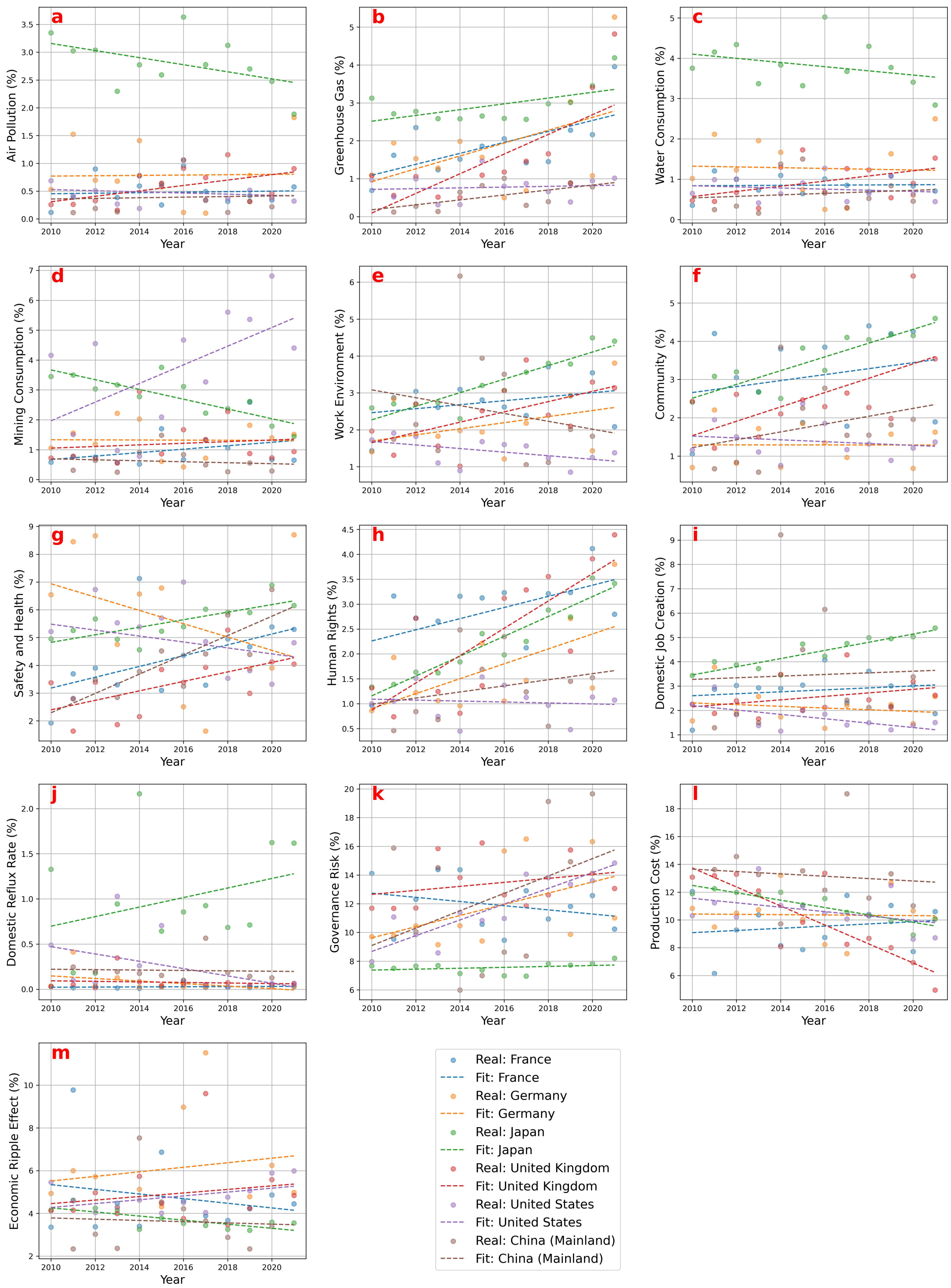

Fig. 6. The ESG Tendency Temporal Variation of Major Countries

Reference:

Aevoae, G. M., Andrieș, A. M., Ongena, S., & Sprincean, N. (2023). ESG and systemic risk. *Applied Economics, 55*(27), 3085-3109.

Alshehhi, A., Nobanee, H., & Khare, N. (2018). The Impact of Sustainability Practices on Corporate Financial Performance: Literature Trends and Future Research Potential. *Sustainability, 10*(2), 494. doi:10.3390/su10020494

Anas, A., Abdul Rashid, H. M., & Annuar, H. A. (2015). The effect of award on CSR disclosures in annual reports of Malaysian PLCs. *Social Responsibility Journal, 11*(4), 831-852. doi:10.1108/SRJ-02-2013-0014

Assaf, C., Monne, J., Harriet, L., & Meunier, L. (2024). ESG investing: Does one score fit all investors' preferences? *Journal of Cleaner Production, 443*, 141094. doi:https://doi.org/10.1016/j.jclepro.2024.141094

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. doi:10.1111/fmii.12132

Bang, J., Ryu, D., & Yu, J. (2023). ESG controversies and investor trading behavior in the Korean market. *Finance Research Letters, 54*, 103750. doi:https://doi.org/10.1016/j.frl.2023.103750

Barko, T., Cremers, M., & Renneboog, L. (2022). Shareholder Engagement on Environmental, Social, and Governance Performance. *Journal of Business Ethics, 180*(2), 777-812. doi:10.1007/s10551-021-04850-z

Baumüller, J., & Sopp, K. (2022). Double materiality and the shift from non-financial to European sustainability reporting: review, outlook and implications. *Journal of Applied Accounting Research, 23*(1), 8-28. doi:10.1108/jaar-04-2021-0114

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

Blitz, D., & Fabozzi, F. J. (2017). Sin stocks revisited: Resolving the sin stock anomaly.

Bodnaruk, A., Loughran, T., & McDonald, B. (2015). Using 10-K Text to Gauge Financial Constraints. *Journal of Financial and Quantitative Analysis, 50*(4), 623-646. doi:10.1017/S0022109015000411

Bolton, P., & Kacperczyk, M. (2021). Do investors care about carbon risk? *Journal of Financial Economics, 142*(2), 517-549. doi:https://doi.org/10.1016/j.jfineco.2021.05.008

Burrell, A. L., Evans, J. P., & De Kauwe, M. G. (2020). Anthropogenic climate change has driven over 5 million km2 of drylands towards desertification. *Nature Communications, 11*(1). doi:10.1038/s41467-020-17710-7

Caglio, A., Melloni, G., & Perego, P. (2020). Informational Content and Assurance of Textual Disclosures: Evidence on Integrated Reporting. *European Accounting Review, 29*(1), 55-83. doi:10.1080/09638180.2019.1677486

Caputo, F., Pizzi, S., Ligorio, L., & Leopizzi, R. (2021). Enhancing environmental information transparency through corporate social responsibility reporting regulation. *Business Strategy and the Environment, 30*(8), 3470-3484. doi:10.1002/bse.2814

Chan, M. C., Watson, J., & Woodliff, D. (2014). Corporate Governance Quality and CSR Disclosures. *Journal of Business Ethics, 125*(1), 59-73. doi:10.1007/s10551-013-1887-8

Chatzitheodorou, K., Skouloudis, A., Evangelinos, K., & Nikolaou, I. (2019). Exploring socially responsible investment perspectives: A literature mapping and an investor classification. *Sustainable Production and Consumption, 19*, 117-129. doi:https://doi.org/10.1016/j.spc.2019.03.006

Chen, Z., & Xie, G. (2022). ESG disclosure and financial performance: Moderating role of ESG investors. *International Review of Financial Analysis, 83*, 102291. doi:https://doi.org/10.1016/j.irfa.2022.102291

Cheng, H., Hu, Y., & Zhao, J. (2009). Meeting China's Water Shortage Crisis: Current Practices and Challenges. *Environmental Science & Technology, 43*(2), 240-244. doi:10.1021/es801934a

Chijoke-Mgbame, A. M., Mgbame, C. O., Akintoye, S., & Ohalehi, P. (2019). The role of corporate governance on CSR disclosure and firm performance in a voluntary environment. *Corporate Governance: The International Journal of Business in Society, 20*(2), 294-306. doi:10.1108/cg-06-2019-0184

Christensen, H. B., Hail, L., & Leuz, C. (2021). Mandatory CSR and sustainability reporting: economic analysis and literature review. *Review of Accounting Studies, 26*(3), 1176-1248. doi:10.1007/s11142-021-09609-5

Cornell, B. (2021). ESG preferences, risk and return. *European Financial Management, 27*(1), 12-19. doi:10.1111/eufm.12295

Cuomo, F., Gaia, S., Girardone, C., & Piserà, S. (2022). The effects of the EU non-financial reporting directive on corporate social responsibility. *The European Journal of Finance*, 1-27. doi:10.1080/1351847x.2022.2113812

D'Augusta, C., Grossetti, F., & Imperatore, C. (2023). Environmental awareness and shareholder proposals: the case of the Deepwater Horizon oil spill disaster. *Corporate Governance: The International Journal of Business in Society*. doi:10.1108/cg-03-2022-0139

Dedoussi, I. C., Eastham, S. D., Monier, E., & Barrett, S. R. H. (2020). Premature mortality related to United States cross-state air pollution. *Nature, 578*(7794), 261-265. doi:10.1038/s41586-020-1983-8

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Di Tommaso, C., & Thornton, J. (2020). Do <scp>ESG</scp> scores effect bank risk taking and value? Evidence from European banks. *Corporate Social Responsibility and Environmental Management, 27*(5), 2286-2298. doi:10.1002/csr.1964

Diener, J. (2023). Impact case or impact washing? An analysis of investors' strategies to influence corporate behavior. *Sustainability Accounting, Management and Policy Journal, 14*(5), 1002-1021.

Dissanayake, D., Tilt, C., & Xydias-Lobo, M. (2016). Sustainability reporting by publicly listed companies in Sri Lanka. *Journal of Cleaner Production, 129*, 169-182. doi:https://doi.org/10.1016/j.jclepro.2016.04.086

Espahbodi, L., Espahbodi, R., Juma, N., & Westbrook, A. (2019). Sustainability priorities, corporate strategy, and investor behavior. *Review of Financial Economics, 37*(1), 149-167.

Feng, Y., Zhu, Q., & Lai, K.-H. (2017). Corporate social responsibility for supply chain management: A literature review and bibliometric analysis. *Journal of Cleaner Production, 158*, 296-307. doi:https://doi.org/10.1016/j.jclepro.2017.05.018

Furman, D., Campisi, J., Verdin, E., Carrera-Bastos, P., Targ, S., Franceschi, C., . . . Slavich, G. M. (2019). Chronic inflammation in the etiology of disease across the life span. *Nature Medicine, 25*(12), 1822-1832. doi:10.1038/s41591-019-0675-0

G20 Green Finance Study Group. (2017). *G20 Green Finance Synthesis Report, 2016*. Retrieved from

Giles, O., & Murphy, D. (2016). SLAPPed: the relationship between SLAPP suits and changed ESG reporting by firms. *Sustainability Accounting, Management and Policy Journal, 7*(1), 44-79. doi:10.1108/SAMPJ-12-2014-0084

Gillan, S. L., Koch, A., & Starks, L. T. (2021). Firms and social responsibility: A review of ESG and CSR research in corporate finance. *Journal of Corporate Finance, 66*, 101889. doi:https://doi.org/10.1016/j.jcorpfin.2021.101889

Giuliani, E. (2016). Human Rights and Corporate Social Responsibility in Developing Countries' Industrial Clusters. *Journal of Business Ethics, 133*(1), 39-54. doi:10.1007/s10551-014-2375-5

Global Sustainable Investment Alliance. (2017). *Global sustainable investment review 2016*. Retrieved from

Goloshchapova, I., Poon, S.-H., Pritchard, M., & Reed, P. (2019). Corporate social responsibility reports: topic analysis and big data approach. *The European Journal of Finance, 25*(17), 1637-1654. doi:10.1080/1351847X.2019.1572637

Gustafsson, M. T., Schilling-Vacaflor, A., & Lenschow, A. (2023). The politics of supply chain regulations: Towards foreign corporate accountability in the area of human rights and the environment? *Regulation & Governance*. doi:10.1111/rego.12526

Gyönyörová, L., Stachoň, M., & Stašek, D. (2023). ESG ratings: relevant information or misleading clue? Evidence from the S&P Global 1200. *Journal of Sustainable Finance & Investment, 13*(2), 1075-1109.

Han, H. H., Lee, J., & Wang, B. (2023). Greenhouse gas emissions, firm value, and the investor base: Evidence from Korea. *Emerging Markets Review, 56*, 101048. doi:https://doi.org/10.1016/j.ememar.2023.101048

Harrison, J. S., Yu, X., & Zhang, Z. (2023). Consistency among common measures of corporate social and sustainability performance. *Journal of Cleaner Production, 391*, 136232. doi:https://doi.org/10.1016/j.jclepro.2023.136232

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. doi:https://doi.org/10.1016/j.jclepro.2023.138960

Hong, H., Li, F. W., & Xu, J. (2019). Climate risks and market efficiency. *Journal of Econometrics, 208*(1), 265-281. doi:https://doi.org/10.1016/j.jeconom.2018.09.015

Huang, A. H., Wang, H., & Yang, Y. (2023). FinBERT: A large language model for extracting information from financial text. *Contemporary Accounting Research, 40*(2), 806-841.

IPCC. (2022). *Climate change 2022: impacts, adaptation and vulnerability*. Retrieved from

Ismail, M. (2009). Corporate social responsibility and its role in community development: An international perspective. *Journal of International social research, 2*(9).

Jiang, L., Gu, Y., & Dai, J. (2023). Environmental, Social, and Governance Taxonomy Simplification: A Hybrid Text Mining Approach. *Journal of Emerging Technologies in Accounting, 20*(1), 305-325.

Keeley, A. R., Chapman, A. J., Yoshida, K., Xie, J., Imbulana, J., Takeda, S., & Managi, S. (2022). ESG metrics and social equity: Investigating commensurability. *Frontiers in Sustainability, 3*, 920955.

Keeley, A. R., Li, C., Takeda, S., Gloria, T., & Managi, S. (2022). The Ultimate Owner of Environmental, Social, and Governance Investment. *Frontiers in Sustainability, 3*. doi:10.3389/frsus.2022.909239

Kiesel, F., & Lücke, F. (2019). ESG in credit ratings and the impact on financial markets. *Financial Markets, Institutions & Instruments, 28*(3), 263-290. doi:10.1111/fmii.12114

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

Landrum, N. E., & Ohsowski, B. (2018). Identifying Worldviews on Corporate Sustainability: A Content Analysis of Corporate Sustainability Reports. *Business Strategy and the Environment, 27*(1), 128-151. doi:10.1002/bse.1989

Lelieveld, J., Evans, J. S., Fnais, M., Giannadaki, D., & Pozzer, A. (2015). The contribution of outdoor air pollution sources to premature mortality on a global scale. *Nature, 525*(7569), 367-371. doi:10.1038/nature15371

Li, Z., Ma, Z., van der Kuijp, T. J., Yuan, Z., & Huang, L. (2014). A review of soil heavy metal pollution from mines in China: Pollution and health risk assessment. *Science of The Total Environment, 468-469*, 843-853. doi:https://doi.org/10.1016/j.scitotenv.2013.08.090

Linsley, P., Abdelbadie, R., & Abdelbadie, R. (2023). The Taskforce on Nature-related Financial Disclosures must engage widely and justify its market-led approach. *Nature Ecology & Evolution*. doi:10.1038/s41559-023-02113-w

Liu, X., Cifuentes-Faura, J., Zhao, S., & Wang, L. (2024). The impact of government environmental attention on firms' ESG performance: Evidence from China. *Research in International Business and Finance, 67*, 102124. doi:https://doi.org/10.1016/j.ribaf.2023.102124

Lokuwaduge, C. S. D. S., & Heenetigala, K. (2017). Integrating Environmental, Social and Governance (ESG) Disclosure for a Sustainable Development: An Australian Study. *Business Strategy and the Environment, 26*(4), 438-450. doi:10.1002/bse.1927

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

Loughran, T., & McDonald, B. (2016). Textual Analysis in Accounting and Finance: A Survey. *Journal of Accounting Research, 54*(4), 1187-1230. doi:10.1111/1475-679x.12123

Loughran, T., McDonald, B., & Otteson, J. R. (2023). How Have Corporate Codes of Ethics Responded to an Era of Increased Scrutiny? *Journal of Business Ethics, 183*(4), 1029-1044. doi:10.1007/s10551-022-05104-2

Loughran, T., McDonald, B., & Yun, H. (2009). A Wolf in Sheep's Clothing: The Use of Ethics-Related Terms in 10-K Reports. *Journal of Business Ethics, 89*(1), 39-49. doi:10.1007/s10551-008-9910-1

LSEG. (2023). Environmental, social and governance scores from LSEG. from LSEG https://www.lseg.com/content/dam/data-analytics/en_us/documents/methodology/lseg-esg-scores-methodology.pdf

Mehra, S., Louka, R., & Zhang, Y. (2022, 2022). *ESGBERT: Language Model to Help with Classification Tasks Related to Companies' Environmental, Social, and Governance Practices*.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems, 26*.

Montero, M. J., Araque, R. A., & Rey, J. M. (2009). Occupational health and safety in the framework of corporate social responsibility. *Safety Science, 47*(10), 1440-1445. doi:https://doi.org/10.1016/j.ssci.2009.03.002

MSCI. (2023a). GICS Sector Definitions. Retrieved from https://www.msci.com/documents/1296102/11185224/GICS+Sector+Definitions+2023.pdf

MSCI. (2023b). The Global Industry Classification Standard (GICS). Retrieved from https://www.msci.com/our-solutions/indexes/gics

MSCI. (2024). ESG Ratings Key Issue Framework. Retrieved from https://www.msci.com/our-solutions/esg-investing/esg-ratings/esg-ratings-key-issue-framework

Nakajima, Y., & Inaba, Y. (2022). Stock market reactions to voluntary integrated reporting. *Journal of Financial Reporting and Accounting, 20*(3/4), 516-541.

Opferkuch, K., Caeiro, S., Salomone, R., & Ramos, T. B. (2021). Circular economy in corporate sustainability reporting: A review of organisational approaches. *Business Strategy and the Environment, 30*(8), 4015-4036. doi:10.1002/bse.2854

Ortiz-Bobea, A., Ault, T. R., Carrillo, C. M., Chambers, R. G., & Lobell, D. B. (2021). Anthropogenic climate change has slowed global agricultural productivity growth. *Nature Climate Change, 11*(4), 306-312. doi:10.1038/s41558-021-01000-1

Oshika, T., & Saka, C. (2017). Sustainability KPIs for integrated reporting. *Social Responsibility Journal, 13*(3), 625-642.

Panthong, S., & Taecharungroj, V. (2021). Which CSR Activities Are Preferred by Local Community Residents? Conjoint and Cluster Analyses. *Sustainability, 13*(19), 10683. doi:10.3390/su131910683

Park, S. R., & Jang, J. Y. (2021). The Impact of ESG Management on Investment Decision: Institutional Investors' Perceptions of Country-Specific ESG Criteria. *International Journal of Financial Studies, 9*(3), 48. doi:10.3390/ijfs9030048

Puri, R., & Catanzaro, B. (2019). Zero-shot text classification with generative language models. *arXiv preprint arXiv:1912.10165*.

Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). Improving language understanding by generative pre-training.

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. *OpenAI blog, 1*(8), 9.

Rahman, H. U., Zahid, M., & Al-Faryan, M. A. S. (2023). ESG and firm performance: The rarely explored moderation of sustainability strategy and top management commitment. *Journal of Cleaner Production, 404*, 136859. doi:https://doi.org/10.1016/j.jclepro.2023.136859

Ramzan, M., Amin, M., & Abbas, M. (2021). How does corporate social responsibility affect financial performance, financial stability, and financial inclusion in the banking sector? Evidence from Pakistan. *Research in International Business and Finance, 55*, 101314. doi:https://doi.org/10.1016/j.ribaf.2020.101314

Raut, R. K., Shastri, N., Mishra, A. K., & Tiwari, A. K. (2023). Investor's values and investment decision towards ESG stocks. *Review of Accounting and Finance*.

Roh, Y., Heo, G., & Whang, S. E. (2021). A Survey on Data Collection for Machine Learning: A Big Data - AI Integration Perspective. *IEEE Transactions on Knowledge and Data Engineering, 33*(4), 1328-1347. doi:10.1109/TKDE.2019.2946162

Rojo-Suárez, J., & Alonso-Conde, A. B. (2024). Have shifts in investor tastes led the market portfolio to capture ESG preferences? *International Review of Financial Analysis, 91*, 103019. doi:https://doi.org/10.1016/j.irfa.2023.103019

Sabbaghi, O. (2022). The impact of news on the volatility of ESG firms. *Global Finance Journal, 51*, 100570. doi:https://doi.org/10.1016/j.gfj.2020.100570

Schilling-Vacaflor, A. (2021). Putting the French Duty of Vigilance Law in Context: Towards Corporate Accountability for Human Rights Violations in the Global South? *Human Rights Review, 22*(1), 109-127. doi:10.1007/s12142-020-00607-9

Schrempf-Stirling, J., & Wettstein, F. (2017). Beyond Guilty Verdicts: Human Rights Litigation and its Impact on Corporations' Human Rights Policies. *Journal of Business Ethics, 145*(3), 545-562. doi:10.1007/s10551-015-2889-5

Schröder, C., & Niekler, A. (2020). A survey of active learning for text classification using deep neural networks. *arXiv preprint arXiv:2008.07267*.

Senadheera, S. S., Withana, P. A., Dissanayake, P. D., Sarkar, B., Chopra, S. S., Rhee, J. H., & Ok, Y. S. (2021). Scoring environment pillar in environmental, social, and governance (ESG) assessment. *Sustainable Environment, 7*(1), 1960097.

Sharfman, M. (1996). The construct validity of the Kinder, Lydenberg & Domini social performance ratings data. *Journal of Business Ethics, 15*(3), 287-296. doi:10.1007/bf00382954

Skouloudis, A., Evangelinos, K., & Kourmousis, F. (2010). Assessing non-financial reports according to the Global Reporting Initiative guidelines: evidence from Greece. *Journal of Cleaner Production, 18*(5), 426-438. doi:https://doi.org/10.1016/j.jclepro.2009.11.015

Sood, K., Pathak, P., Jain, J., & Gupta, S. (2023). How does an investor prioritize ESG factors in India? An assessment based on fuzzy AHP. *Managerial Finance, 49*(1), 66-87.

Stanton, P., & Stanton, J. (2002). Corporate annual reports: research perspectives used. *Accounting, Auditing & Accountability Journal, 15*(4), 478-500. doi:10.1108/09513570210440568

Sundaram, A. K., & Inkpen, A. C. (2004). The Corporate Objective Revisited. *Organization Science, 15*(3), 350-363. doi:10.1287/orsc.1040.0068

Tang, H., Xiong, L., & Peng, R. (2024). The mediating role of investor confidence on ESG performance and firm value: Evidence from Chinese listed firms. *Finance Research Letters, 61*, 104988. doi:https://doi.org/10.1016/j.frl.2024.104988

Tilling, M. V., & Tilt, C. A. (2010). The edge of legitimacy. *Accounting, Auditing & Accountability Journal, 23*(1), 55-81. doi:10.1108/09513571011010600

Tomar, S. (2023). Greenhouse Gas Disclosure and Emissions Benchmarking. *Journal of Accounting Research, 61*(2), 451-492. doi:10.1111/1475-679x.12473

Van Duuren, E., Plantinga, A., & Scholtens, B. (2016). ESG Integration and the Investment Management Process: Fundamental Investing Reinvented. *Journal of Business Ethics, 138*(3), 525-533. doi:10.1007/s10551-015-2610-8

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Aidan, . . . Polosukhin, I. (2017). Attention Is All You Need. *arXiv pre-print server*. doi:None

arxiv:1706.03762

Wan, G., Dawod, A. Y., Chanaim, S., & Ramasamy, S. S. (2023). Hotspots and trends of environmental, social and governance (ESG) research: a bibliometric analysis. *Data Science and Management, 6*(2), 65-75. doi:https://doi.org/10.1016/j.dsm.2023.03.001

Wu, S., Palm-Forster, L. H., & Messer, K. D. (2021). Impact of peer comparisons and firm heterogeneity on nonpoint source water pollution: An experimental study. *Resource and Energy Economics, 63*, 101142. doi:https://doi.org/10.1016/j.reseneeco.2019.101142

Xie, J., Nozawa, W., Yagi, M., Fujii, H., & Managi, S. (2019). Do environmental, social, and governance activities improve corporate financial performance? *Business Strategy and the Environment, 28*(2), 286-300. doi:10.1002/bse.2224

Zhang, X., Wang, Z., Zhong, X., Yang, S., & Siddik, A. B. (2022). Do Green Banking Activities Improve the Banks' Environmental Performance? The Mediating Effect of Green Financing. *Sustainability, 14*(2), 989. doi:10.3390/su14020989

Zhang, Z., Su, Z., Wang, K., & Zhang, Y. (2022). Corporate environmental information disclosure and stock price crash risk: Evidence from Chinese listed heavily polluting companies. *Energy Economics, 112*, 106116. doi:https://doi.org/10.1016/j.eneco.2022.106116

Zhou, G., Liu, L., & Luo, S. (2022). Sustainable development, ESG performance and company market value: Mediating effect of financial performance. *Business Strategy and the Environment, 31*(7), 3371-3387. doi:10.1002/bse.3089