Renovation in Environmental, Social, and Governance (ESG) Research: The Application of Machine Learning

Article *in* Asian Review of Accounting · September 2024

DOI: 10.1108/ARA-07-2023-0201

Joseph Zhang
Bowling Green State University

Renovation in environmental, social and governance (ESG) research: the application of machine learning

Abby Yaqing Zhang and Joseph H. Zhang
School of Accountancy, The University of Memphis, Memphis, Tennessee, USA

Received 30 July 2023
Revised 18 September 2023
25 September 2023
Accepted 7 October 2023

Abstract

Purpose – Environmental, social and governance (ESG) factors have become increasingly important in investment decisions, leading to a surge in ESG investing and the rise of sustainable investment assets. Nevertheless, challenges in ESG disclosure, such as quantifying unstructured data, lack of guidelines and comparability, rampantly exist. ESG rating agencies play a crucial role in assessing corporate ESG performance, but concerns over their credibility and reliability persist. To address these issues, researchers are increasingly utilizing machine learning (ML) tools to enhance ESG reporting and evaluation. By leveraging ML, accounting practitioners and researchers gain deeper insights into the relationship between ESG practices and financial performance, offering a more data-driven understanding of ESG impacts on business communities.

Design/methodology/approach – The authors review the current research on ESG disclosure and ESG performance disagreement, followed by the review of current ESG research with ML tools in three areas: connecting ML with ESG disclosures, integrating ML with ESG rating disagreement and employing ML with ESG in other settings. By comparing different research's ML applications in ESG research, the authors conclude the positive and negative sides of those research studies.

Findings – The practice of ESG reporting and assurance is on the rise, but still in its technical infancy. ML methods offer advantages over traditional approaches in accounting, efficiently handling large, unstructured data and capturing complex patterns, contributing to their superiority. ML methods excel in prediction accuracy, making them ideal for tasks like fraud detection and financial forecasting. Their adaptability and feature interaction capabilities make them well-suited for addressing diverse and evolving accounting problems, surpassing traditional methods in accuracy and insight.

Originality/value – The authors broadly review the accounting research with the ML method in ESG-related issues. By emphasizing the advantages of ML compared to traditional methods, the authors offer suggestions for future research in ML applications in ESG-related fields.

Keywords Environmental, Social, Machine learning, ESG reporting and governance (ESG), ESG rating disagreement

Paper type General review

1. Introduction

Environmental, social and governance (ESG) is a stakeholder-centric approach that advocates considering ESG factors when deciding to invest in companies. The term "ESG" emerged in 2004 with the release of the report "Who Cares Wins" by the United Nations Global Compact (UNGC) [1]. There has been a notable surge in investor interest in ESG investing in recent years. The growing interest in ESG among capital market participants and regulators has spurred an expansion of academic research in the accounting literature that explores ESG issues. We organize the literature on ESG issues into two categories: ESG disclosure research and ESG performance evaluation (i.e. reports of ESG rating agencies). Initial studies identify the specific characteristics of firms that are associated with ESG disclosures. Cowen *et al.* (1987) find that corporate size and industry category have significant effects on firms' ESG disclosure choices, followed by ownership structure (Khlif *et al.*, 2017; McGuinness *et al.*, 2017; Abeysekera and Fernando, 2020), corporate governance (Bui *et al.*, 2020; Guo and Yu, 2022) and leadership

Asian Review of Accounting

1321-7348
DOI 10.1108/ARA-07-2023-0201

characteristics (Barnea and Rubin, 2010; Borghesi *et al.*, 2014; Hegde and Mishra, 2019; Iliev and Roth, 2023). Further, country-level determinants of ESG disclosures have been explored (Wanderley *et al.*, 2008; Baldini *et al.*, 2018).

A myriad of theories also explain the determinants of ESG disclosures, such as stakeholder theory, e.g. pressure from stakeholders (Kim *et al.*, 2012; Ioannou *et al.*, 2016; Martin and Moser, 2016; Naughton *et al.*, 2019), legitimacy theory, e.g. government requirement (Michelon *et al.*, 2020; Pinnuck *et al.*, 2021), agency theory (Christensen *et al.*, 2022) and signaling theory (Ryou *et al.*, 2022). The consequences of ESG disclosures have been explored widely, too. Many studies have confirmed the effects of ESG on firm performance (Brown *et al.*, 2006; Rastogi *et al.*, 2023), investors (Naughton *et al.*, 2019), analysts (Dhaliwal *et al.*, 2012) and other stakeholders. For example, Brown *et al.* (2006) indicate that firm value increases with firms' philanthropic engagement. The association between ESG disclosures and firms' financial distress probability (Al-Hadi *et al.*, 2017), corporate value (Lougee and Wallace, 2008), competitive advantages (Porter and Kramer, 2006) and other firm value aspects have been explored as well.

In the meantime, studies have shown that companies that disclose ESG information experience a variety of economically significant effects, including changes in their cost of capital, stock price and access to capital. For example, high-quality ESG disclosures are associated with better access to finance, i.e. lower capital constraints (Cheng *et al.*, 2014), lower analyst forecast errors (Byard *et al.*, 2011; Dhaliwal *et al.*, 2012), fewer earnings management (Kim *et al.*, 2012; Gao and Zhang, 2015; Dang *et al.*, 2021), reduced tax payments (Davis *et al.*, 2016) and fewer audit fees (Burke *et al.*, 2019).

While numerous papers apply ESG scores from rating agencies to evaluate firms' ESG disclosure quality, and these third-party raters strive to establish consistent and reliable metrics for evaluating firms' ESG performance, skeptics doubt the credibility and reliability of the raters' rating due to their variable data collections process, such as collecting firms' ESG data through firms' SEC filings, press releases and social media, capturing information from government, regulatory and non-governmental organizations and getting private information through solicited questionnaires to firms. Earlier evidence, such as Chatterji *et al.* (2016), reveals a notable lack of agreement among six prominent rating agencies, raising significant concerns about the usefulness and reliability of these ratings. Further, Berg *et al.* (2022) point out that the divergence observed in ESG ratings can be attributed to how different rating agencies measure, define and assign weights to their ESG ratings. The mixed results between firms' ESG disclosures and rating agencies' disagreements exacerbate the doubts about the credibility of rating agencies' ESG evaluations, where Christensen *et al.* (2022) find that increased ESG disclosures amplify the level of disagreement observed among ESG ratings. In contrast, Kimbrough *et al.* (2022) argue that firms that voluntarily issue ESG reports experience lower levels of disagreement among ESG rating agencies.

Machine learning (ML), a subfield of artificial intelligence (AI), has gained significant attention in accounting research due to its potential to enhance decision-making processes, improve financial analysis, detect fraudulent activities and conduct textual analysis and topical modeling. A growing number of accounting researchers have applied ML in accounting studies. For example, researchers use ML to predict firms' financial fraud behavior (Perols, 2011; Bao *et al.*, 2020), bankruptcy (Barboza *et al.*, 2017), misstatements (Bertomeu *et al.*, 2021) and tax avoidance (Guenther *et al.*, 2023). Another trendy ML application in accounting research focuses on textual analysis of firms' disclosures. An early study by Li (2010) introduces the naïve Bayesian algorithm, a supervised ML algorithm used for classification tasks in the Management Discussion and Analysis (MD&A) setting.

Researchers also focus on the modal word frequency and sentiment analysis of corporate disclosure tone by examining earnings conference calls (Li *et al.*, 2021), mandatory SEC filings (Huang *et al.*, 2023), press releases (Henry and Leone, 2016), analyst reports (Huang *et al.*, 2014), critical audit matters (Liu *et al.*, 2022) and social media (Booker *et al.*, 2023). While many articles have explored the ML application in various firms' disclosures, the ML application in ESG disclosures, known as nonfinancial disclosure, has less evidence.

In short, the practice of ESG reporting and assurance is on the rise, but still in its technical infancy. We discuss the application of wide-spectrum ML tools in ESG-related research in three areas: connecting ML with ESG disclosures, integrating ML with ESG rating disagreement and employing ML with ESG in other settings, like corporate ESG issues mentioned in conference calls, or financial reports, or social media.

The remainder of the paper is organized as follows. Section 2 reviews the current research on ESG disclosure and ESG performance disagreement. Section 3 discusses the current ESG research with ML tools. Section 4 concludes. We discuss more ESG research and regulations in Appendix.

2. Research on ESG disclosure and ESG performance evaluation disagreements

In response to the increased interest and demand from investors, regulators and other stakeholders, many companies voluntarily or are required to provide ESG disclosures using a variety of reporting standards. Academic research seeks to find the determinants pressing firms to provide ESG disclosures and investigate the economic consequences of the firms' ESG disclosures. The initial evidence indicates the benefits of firms' ESG disclosures. Yet, the rating agencies' disagreements have raised concerns about the credibility of the firms' ESG performance evaluation. This section briefly reviews the extant research on ESG disclosure in two primary areas: (1) determinants and consequences and (2) ESG performance evaluation disagreement.

2.1 Research on ESG disclosure

Increasingly, firms are engaging in ESG programs to meet the growing interest of investors and other stakeholders in ESG issues. Following the steps, academic researchers have paid attention to ESG performance and reporting as well. We first review the determinants that drive firms to disclose ESG reports. Then, we discuss the consequences of the ESG disclosures.

2.1.1 Determinants of ESG disclosures. Cowen *et al.* (1987), one of the antecedents, find that corporate size has a significant impact on firms' specific disclosures, like disclosure of energy matters. More researchers seek to explore other firm characteristics, promoting firms' movements to ESG disclosures (McGuinness *et al.*, 2017; Abeysekera and Fernando, 2020; Bui *et al.*, 2020). One research underscores the significant role of qualified foreign institutional investors in establishing a competitive advantage through ESG engagements and adds to the literature on leadership characteristics by demonstrating a positive correlation between greater gender representation in top management and improved ESG performance (McGuinness *et al.*, 2017). In the setting of family firms, Abeysekera and Fernando (2020) document that family-owned businesses demonstrate a higher responsibility towards shareholders when making environmental investments.

In a similar vein, Bui *et al.* (2020) explore the effects of ESG strategies on firms' ESG-related disclosures and find that climate governance demonstrates a connection to the alignment between carbon disclosure and carbon performance. Further, board directors' experience in ESG issues (Iliev and Roth, 2023), CEO age, gender, political affiliation contributions (Borghesi *et al.*, 2014) and CEOs' marriage (Hegde and Mishra, 2019) have been validated in the association between leadership characteristics and firms' ESG engagement.

Other studies have investigated how country-level factors can affect firms' ESG disclosures. Baldini *et al.* (2018) use a cross-country sample to probe the country-level factors, pressing companies to release ESG disclosures and find that the political system (legal framework and corruption), labor system (labor protection and unemployment rate) and cultural system (social cohesion and equal opportunities) have a substantial influence on the firms' ESG disclosures practices. Similarly, Wanderley *et al.* (2008) find evidence to support the notion that country-of-origin factors significantly affect firms' ESG disclosures on the corporate websites in emerging countries.

2.1.2 Consequences of ESG disclosures. The consequences of ESG disclosures have been explored widely. Researchers examine the effects of ESG disclosure on capital market participants and stakeholders – how such disclosure influences their behavior and decision-making. Initial evidence investigates the effects of ESG disclosures on firm performance. However, the results regarding the association between ESG disclosures and firm performance are mixed. Some studies suggest that the association is positive. That is to say, ESG disclosures benefit firm profitability by positively influencing a firm's relationship with key stakeholders (Brown *et al.*, 2006; Naughton *et al.*, 2019). Specifically, Brown *et al.* (2006) show the earlier evidence that firm value increases with firms' engagement in philanthropy. In contrast, other researchers suggest a negative association between ESG disclosures and the value that investors assign to those firms in the extent of emission-produced setting (Clarkson *et al.*, 2015; Griffin *et al.*, 2017).

Further, a company can gain the most significant competitive advantage by identifying the specific ESG issues that it is best suited to address (Porter and Kramer, 2006) and mitigate the firm's financial distress, especially for firms in mature life cycle stages (Al-Hadi *et al.*, 2017), fosters a sense of pride, enhances recruitment efforts, aids in retaining top talent and boosts employee productivity (Lougee and Wallace, 2008).

In addition, studies on the consequences of ESG disclosure on the capital market have revealed that firms engaging in ESG information disclosure can experience a range of economically significant effects. Firms with outstanding ESG performance have been confirmed to have better access to finance, which can be attributed to the decrease in agency costs arising from improved stakeholder engagement and the decrease in informational asymmetry due to enhanced transparency (Cheng *et al.*, 2014). To answer the question of whether firms with superior ESG performance exhibit distinct behaviors in their financial reporting compared to other firms, Kim *et al.* (2012) demonstrate that firms with higher-quality ESG performance have fewer earnings management, manipulations in real operating activities and SEC investigations. Similar trends for firms with good ESG performance can be found in fewer tax payments (Davis *et al.*, 2016) and audit fees (Burke *et al.*, 2019).

Lastly, increasing research has paid attention to the effects of ESG disclosures on other stakeholders, like customers and employees. Baron (2008) provides that firms with better ESG performance may attract customers who highly value sustainable practices and are willing to pay a premium for them. However, he also indicates that the findings are not definitive. Different factors and contexts could influence the relationship. Recently, researchers (e.g. Dube and Zhu, 2021, Lalova, 2023, Welch and Yoon, 2023) use Glassdoor review data to investigate how firms' ESG disclosure practices affect employees.

2.2 Research on ESG performance evaluation disagreement

The growing interest of capital market participants in firms' ESG disclosure practices has led to the development of third-party ESG performance evaluation rating agencies. These agencies provide information to market participants, such as investors, analysts and corporate managers, about the quality of firms' ESG programs and potential risks posed by

ESG issues. Investors depend on the information to make informed investment decisions, while corporations utilize ratings to obtain external feedback regarding the effectiveness and quality of their sustainability initiatives. Prior research uses rating agencies' ratings as the method to measure firms' ESG quality and seeks to find the associations between ESG quality and financial and nonfinancial outcomes (Chatterji and Toffel, 2010; Grewal *et al.*, 2019; Zhao and Huang, 2021; Serafeim and Yoon, 2023; Welch and Yoon, 2023).

However, due to the ratings' multidimensionality (e.g. the materiality of separate E, S, G dimension varies across industries, even across companies), lack of completeness (i.e. rating agencies have to make decisions on handling missing data, including the choice to omit the data point or make assumptions to allow for comparisons across companies in dealing with non-publicly reported information), lack of standardization (e.g. companies report information using different scales, such as raw numbers, time scales, or percentage scales, making direct comparisons of variables challenging) and lack of consistency (i.e. rating providers may modify historical data in models to enhance performance by incorporating new or improved information, but this can inadvertently make the model appear more predictive than its original state), skeptics doubt the credibility and reliability of the raters' ratings. This matter holds significant importance as the lack of consensus on defining good ESG performance can potentially mislead market participants who rely on ESG ratings. As a result, an emerging stream of literature has addressed the issue of significant discrepancies in ratings provided by different raters for the same company.

Chatterji *et al.* (2016), the pioneer in addressing the ESG rating disagreement issues, document a disagreement among six well-established raters, namely, *KLD*, *Asset4*, *Innovest*, *DJSI*, *FTSE4Good* and *Calvert*, leading to widespread criticism regarding the usefulness and reliability of these ratings. They provide two terms to explain the rating divergence: "theorization" and "commensurability," in other words, what metrics raters assess and the significance of their evaluations, and whether ratings enable comparisons across different raters. However, their empirical results do not provide a clear understanding of the extent to which each of these components contributes to the observed divergence. Berg *et al.* (2022) fill up this gap by identifying three divergence dimensions: scope (the range of attributes that rating providers aim to measure in their evaluations), measurement (the specific measures or indicators used by rating agencies to assess the same attributes) and weighting (assigning relative importance or weights to different attributes in the overall rating calculation).

In addition, comparing the ESG data of two different companies is a challenging task due to the absence of standardized guidelines governing the reporting of ESG information (Amel-Zadeh and Serafeim, 2018). Hence, the research seeks to answer whether firms' ESG disclosure can explain some of this disagreement. However, the mixed results exacerbate doubts about the credibility of ESG evaluations by rating agencies. For example, Christensen *et al.* (2022) find that increased ESG disclosures amplify the level of disagreement observed among ESG ratings, while Kimbrough *et al.* (2022) argue that firms that voluntarily issue ESG reports experience lower levels of disagreement among ESG rating agencies.

There is limited literature on the effects of ESG rating divergence. Gibson Brandon *et al.* (2021) examine the market response to ESG rating disagreement by analyzing the monthly returns associated with the monthly ESG rating disagreement and find a positive relationship between stock returns and ESG rating disagreement, implying that firms with high ESG rating disagreement are shown to have a higher risk premium. As an extension to Gibson Brandon *et al.* (2021), the study by Serafeim and Yoon (2023) reveals that the consensus rating has predictive power for future news. However, its effectiveness diminishes for firms with significant disagreement among raters. Furthermore, the impact of news on market reaction is significantly influenced by the consensus rating.

3. Research on ESG with ML tools

Given the rating disagreements issue among rating agencies that serve as the information intermediaries for ESG information in financial markets and have influences on stakeholders' decision-making, the demand for more accurate, efficient and effective methods to evaluate firms' ESG practices cannot be ignored. Compared to traditional statistical methods, ML methods are recognized for their ability to efficiently leverage large volumes of data. More recent studies have resorted to ML tools to improve the assessment of ESG quality. ML, a subfield of artificial intelligence (AI), is increasingly being used in accounting research to improve decision-making, financial analysis, fraud detection and textual analysis. ML algorithms can learn from large amounts of data and identify patterns that would be difficult for humans to see. This makes them well-suited for tasks such as predicting bankruptcy, detecting fraud and analyzing financial reporting and nonfinancial disclosures.

A growing number of accounting researchers have applied ML in accounting studies. For example, researchers use ML to predict firms' financial fraud behavior (Perols, 2011; Bao *et al.*, 2020), bankruptcy (Barboza *et al.*, 2017), misstatements (Bertomeu *et al.*, 2021), effective tax rates (Guenther *et al.*, 2023), auditor switches (Hunt *et al.*, 2021) and selecting directors (Erel *et al.*, 2021). Another trend of ML in accounting research focuses on textual analysis of corporate disclosures (Li, 2010; Huang *et al.*, 2014, 2023; Henry and Leone, 2016; Li *et al.*, 2021; Liu *et al.*, 2022; Booker *et al.*, 2023). In the ESG setting, several researchers simply measure firms' ESG practices by an indicator variable to capture whether a firm has ESG disclosure (Dhaliwal *et al.*, 2012; Hoi *et al.*, 2013; Dai *et al.*, 2023). However, ESG disclosures, mostly non-numeric information disclosures, have a great deal of unconstructed data. The application of ML can quantify firms' ESG conversations, helping business communities objectively assess the effects of firms' ESG practices on business operations.

Specifically, on the one hand, connecting ML with ESG practices could be more powerful because of ML's ability to handle unstructured data from firms' non-numeric information. On the other hand, connecting ML with firms' ESG practices directly can let firms "speak for themselves" instead of relying on ESG rating scores, especially the concerns of rating disagreement. Nonetheless, the opaqueness of the process between input and output in ML algorithms, often rendering them as black boxes, poses a barrier to the adoption of ML models. Therefore, researchers need to thoroughly evaluate the balance between predictive accuracy and interpretability, as well as consider data quality and features, when choosing and deploying ML models in accounting research on ESG.

Corporate ESG reports' non-numeric and textual features allow researchers to use natural language processing (NLP), an ML technology that empowers computers to interpret, manipulate and understand human language, to analyze and the analysis results can reduce information asymmetry and enhance market efficiency and resource allocations. Popular NLP models, frequently employed in the finance and accounting fields, include word embedding models, topic models and sentiment analysis models. Word embedding models generate compact vector representations of words in a continuous space, capturing semantic relationships between words and facilitating analysis of financial news sentiment and word/text similarity identification in financial reports. Word2Vec and BERT (bidirectional encoder representations from transformers) are two widely used word embedding models in finance and accounting research. Word2Vec uses neural networks to learn word embeddings from large text corpora. BERT, another common word embedding model, is a deep learning NLP model that undergoes training on a vast collection of text documents and utilizes encoding to represent text mathematically while considering the contextual information surrounding each word.

A topic model, such as Latent Dirichlet allocation (LDA), is an ML technique that automatically identifies main themes or latent patterns within a document collection, grouping-related words to form coherent topics. Sentiment analysis model, an NLP task also referred to as opinion mining, aims to classify text data into various sentiment categories, such as positive, negative, neutral, or more nuanced emotions like happy, sad and angry, using specialized ML models. Researchers analyze ESG-related issues using NLP, collect related textual data from sources like news articles, company reports, social media and regulatory filings and preprocess the data by eliminating noise, tokenizing the text and converting it to an appropriate format for NLP analysis. After cleaning data, several NLP models, such as word embedding models, topic models and sentiment analysis models, are used to analyze textual data and extract insights from corporate ESG practices. Freshly, large language models (LLM), a deep-learning-based NLP algorithm with numerous parameters, such as OpenAI GPT, have been introduced in ESG-related issues, allowing researchers to learn semantic and syntactic relationships from extensive text data and consider context during text summarization. Moreover, other ML tools, like tree-based models (e.g. decision trees and random forest, tree-based models are supervised ML algorithms used for tasks such as classification and regression), are also applied in ESG-related issues.

To the best of our knowledge, a study by Huang *et al.* (2023) is the first and only one connecting ESG with ML among the top five accounting journals. It is obvious that there is limited evidence in accounting research to explore the ML application in ESG settings and a broad future for accounting researchers to investigate ESG issues with ML tools. We review the recent studies, not exhaustive, with the application of ML in ESG research and categorize them into three areas: ML in ESG-related disclosure, ML in ESG ratings and ML in other settings, like ESG-related issues from earnings conference call transcripts or firms' financial reports.

3.1 ML in ESG-related disclosures

Textual analysis, also known as text analysis or text mining, is the process of examining and extracting meaningful information from written or textual data. It involves applying various computational and analytical techniques to understand the patterns, structures and relationships within a text. ESG disclosure, known as a nonfinancial (often non-numeric) information disclosure, comprises both numerical or quantitative data and unstructured textual or qualitative information. Prior research has shown stock market reaction to ESG disclosures (Brown *et al.*, 2006; Dhaliwal *et al.*, 2012; Naughton *et al.*, 2019). However, a general description of ESG disclosures, such as whether to have ESG disclosure or to release ESG reports with fewer pages, is insufficient to predict the market reaction. As a result, textual analysis of ESG disclosures has become an additional tool that aids in reducing information asymmetry and enhancing market efficiency.

The research by Huang *et al.* (2023), the first one connecting ESG with ML among the top five accounting journals, manually categorizes 2,000 sentences from firm disclosures into environmental, social, governance and non-ESG categories following the MSCI ESG rating methodology [2]. Comparing FinBERT, a state-of-the-art large language model designed for financial texts based on Google's BERT algorithm [3], with other ML models and the dictionary approach (Loughran and McDonald, 2016), FinBERT achieves higher accuracy in sentiment classification for labeled ESG sentences. However, their study's limitation lies in using only 2,000 manually labeled ESG sentences, reducing the generalizability and validation of the findings connecting ESG with ML. Nonetheless, the paper represents an initial step in connecting ESG with ML by incorporating contextual information in the accounting field.

Some researchers seek to use ML to identify climate-related risks from firms' disclosures. Extension to TCFD (2022), which has conducted an "AI review" based on BERT architecture to identify compliance with the TCFD Recommended Disclosures without assessing the informativeness of disclosed information, Luccioni *et al.* (2020) develop ClimateQA, a custom transformer-based model, using NLP advancements to identify climate-relevant sections in financial reports through question-answering [4]. Similarly, Bingler *et al.* (2022) introduce ClimateBERT, a fine-tuned BERT model with text mining algorithms, analyzing climate-risk disclosures across TCFD's main categories. They find TCFD recommendations significantly impact TCFD-supporting companies' disclosures, yet strategy, metrics and targets disclosure lags, indicating potential issues with voluntary commitments and selective non-material disclosure.

Finally, Sautner *et al.* (2023) modify King *et al.*'s (2017) ML keyword discovery algorithm, generating climate change bigrams by selecting expert-recognized terms, capturing climate change exposures like opportunity, physical impacts and regulatory shocks in corporate conference calls. While all these studies on climate issues have focused on analyzing the frequency of climate-related disclosures, the quality and materiality of these disclosures have been inadequately addressed and remain unclear. In addition, previous studies primarily employ sentence-level classification. Friederich *et al.* (2021) recognize the necessity for additional context when disclosing risks and refine the ML technique by DistilBERT [5] and RoBERTa Large [6] to a paragraph-level classification based on the 50 largest publicly traded companies. Del Vitto *et al.* (2023) also use the Refinitiv data and construct algorithms to predict companies' individual E, S and G sustainability ratings. Their research shows that ML models are able to predict Refinitiv ESG scores if trained with a suitable selection of features.

Most of the aforementioned studies employ the word embedding model when they analyze the ESG reports issues. Some researchers (e.g. Chae and Park, 2018; Goloshchapova *et al.*, 2019; Kiesel and Lücke, 2019) pay attention to the topic models to extract information from corporate ESG practice. LDA is a probabilistic generative model used in NLP and ML for topic modeling, where it assumes that documents with similar topics share common words, enabling the essence of each document to be summarized and represented by its highest-weighted topic composition. Specifically, Kiesel and Lücke (2019) utilize the LDA model to identify ESG topics in 3,719 Moody's credit rating reports and reveal a clear but modest influence of ESG performance on rating decisions. Similarly, using LDA, Goloshchapova *et al.* (2019) analyze publicly available ESG reports of firms in MSCI Europe, revealing shared topics, including "employees safety", "employees training support", "carbon emission", "human right", "efficient power" and "healthcare medicines", in the reports of publicly listed companies in Europe and the UK.

Recently, Föhr *et al.* (2023) introduce ChatGPT, a foundation model from OpenAI. ChatGPT is an AI model that can be used for a variety of tasks, including text, image and audio generation. It is also capable of transfer learning and zero-shot learning, which means that it can apply knowledge from one task to a new, related task without being fine-tuned. This makes ChatGPT well-suited for assisting auditors in complying with the EU Taxonomies and providing essential insights during the auditing process. The application of ChatGPT in identifying ESG-related issues during the auditing process is a significant innovation. However, it is imperative to acknowledge that foundation models, like ChatGPT, may sometimes draw incorrect inferences or misrepresent data due to their inherent limitations. As Föhr *et al.* (2023) suggest, the promising collaboration of foundation models and human expertise could improve the auditing process, especially with the maturity of foundation models.

3.2 ML in ESG ratings

Companies that adopt strong ESG practices appear to gain a competitive edge from various avenues, resulting in a more effective allocation of resources. Market participants use firms' ESG scores (assessment from ESG raters) to consider and incorporate these factors into their decision-making processes. However, ESG scores of individual firms exhibit significant diversity when assessed by different agencies, leading to varying evaluations of their ESG performance. Limited literature addresses this inconsistency issue among ESG ratings by ML tools. Lanza *et al.* (2020) introduce an innovative method to resolve the prevailing disparities in ESG ratings by leveraging ML to pinpoint the ESG indicators that significantly contribute to constructing efficient portfolios. Obtaining ESG metrics from Refinitiv-Asset 4 [7] and MSCI [8], they apply a tree-based ML approach to select the ten most relevant ESG factors among 220 ESG metrics (omitting some overlapping indicators) and combine those indicators to construct portfolios [9]. Among the most material ESG indicators (17 indicators), half of them are linked to environmental matters (especially transition risk, referring to companies' exposure and capability to handle climate change risk), highlighting the crucial role of environmental issues in influencing the performance of the equity portfolio. However, their paper only adopts two ESG raters, which cover European companies without US companies.

Aiba *et al.* (2019) seek to find the comparability between ESG raters' assessments and GRI indicators by adopting two methods to build comparability: word-level vectorization (segmenting the text into individual words, converting each word into a vector, weighting the word vectors based on importance, and then summing them to create a vector representing the entire piece of text) and sentence-level vectorization (vectorizing entire sentences in a text and subsequently summing these sentence vectors together). By employing two text mining tools, namely, term frequency-inverse document frequency (TF-IDF) algorithm and BERT in separate vectorization, they develop two similarity metrics, utilizing them to align ESG raters' assessments with GRI indicators, thereby calculating topic weights for MSCI and FTSE scores. Their successful approach identifies the most valued GRI indicators by ESG raters and strategically prioritizes disclosure areas in each sample company, but the lack of transparency in rating agencies' algorithms and the significant variation in ESG scores from different agencies make creating a taxonomy based on matching GRI indicators and ESG scores less advantageous.

Further, some researchers have attempted to enhance the comparability of ESG disclosures by ML tools. Different from Aiba *et al.* (2019), Jiang *et al.* (2023) align GRI indicators with ESG reports to create a more straightforward classification system instead of correlating ESG ratings. Combining text mining techniques for efficient processing and manual human judgment for more accurate categorization, the hybrid approach seeks to strike a harmonious balance between automation and human expertise. Efforts to make ESG taxonomies comparable are advancing in their research, whereas their study only relies on GRI indicators, potentially reducing the generalizability, credibility and comparability of ESG reports.

3.3 ML in other settings with respect to ESG practice

Other research focuses on connecting ML with ESG practices in other settings, not limited to ESG disclosures and ESG raters' divergence. For example, rooted in Glassdoor employee reviews, the paper by Briscoe-Tran (2023) uses a word-embedding model to develop an internal perspective of corporate ESG practices. To overcome the challenge of creating comprehensive dictionaries for ESG topics tailored to employee reviews, the authors generate seed word lists for each ESG category by extracting commonly used words and phrases related to E, S and G issues from ESG rating methodologies and academic articles, and then they employ an ML model, namely, Word2Vec, trained on 10.4 million employee reviews, to grasp the semantics of words [10].

The results demonstrate that employees' discussion of ESG topics offers significant information about firms' ESG practices, enabling predictions of various indicators of a firm's future ESG-related outcomes, particularly in the S and G dimensions, surpassing existing

ESG ratings. However, employees and rating agencies may have divergent priorities regarding ESG topics, leading to differences in their assessments and concerns and biasing the analysis results.

Likewise, leveraging a word embedding model, the research conducted by Lin *et al.* (2023) examines more than 210,000 annual reports from 24,271 public firms in thirty (30) countries/regions spanning 2001 to 2020. Their results demonstrate rising trends in E&S (environmental and social) disclosure, including increased length, boilerplate language usage, stickiness and infographics, while specificity decreases. Li *et al.* (2021) apply Word2Vec in analyzing firms' G dimension of ESG from earnings conference call transcripts, developing a comprehensive culture dictionary to assess firms' corporate cultural values, encompassing innovation, integrity, quality, respect and teamwork.

In addition, Klusak *et al.* (2023) build a climate-adjusted sovereign credit rating. In this study, a random forest ML model is developed and trained on sovereign credit ratings for 108 countries to predict future sovereign credit ratings accurately. Subsequently, climate economic models and S&P Global Rating's own natural disaster risk assessments are integrated to develop climate-adjusted macroeconomic data, which is then used in the rating prediction model under different policy and warming scenarios to forecast the influence of climate change on sovereign ratings for future years 2030, 2050, 2070 and 2100. The study by Klusak *et al.* (2023) is the pioneer in simulating the impact of future climate change on sovereign credit ratings, driving future research to pay more attention to the effects of ESG practices on financial performance. Jain *et al.* (2023) explore how GPT 3.5 responds to ESG questions, utilizing APIs like Alpaca [11], News API [12] and OpenAI, [13], to collect and analyze Tesla stock returns and news sentiment. Through response analysis, their study's findings reveal that GPT 3.5 can provide informative and precise responses to ESG-related prompts.

It has been witnessed that the versatile power of ML enhances ESG reporting by automating data collection and analysis, enabling trend identification, correlation discovery and predictive analytics, resulting in more accurate and efficient reporting. Integrating ML tools with ESG practices has become prominent in academic research, addressing ESG-related disclosures, rating discrepancies, practice comparability, internal ESG performance assessment (e.g. employees' perspective), annual reports, social media, and more. Compared to the traditional logistic method in accounting research, which emphasizes causal inference by maximizing accuracy through minimizing bias arising from model misspecification and obtaining a highly precise representation of the underlying theory or enforcing a structure predetermined on the relationships between variables, ML methods outperform in the abilities to efficiently leverage and handle with large volumes of unstructured data, capture intricate patterns and non-linear relationships while imposing fewer limiting assumptions, attain superior prediction accuracy and classification performance, automatically acquire feature interaction and identify relevant variables.

The current ESG-related research with ML tools mainly applies word embedding models due to their excellent ability to learn context-specific meanings of words and phrases in large-scale datasets. The topic models are also considered by researchers because latent themes or topics within a collection of documents, such as corporate reports, news articles and social media posts, can be identified through ML tools. Meanwhile, the emergence of OpenAI's ChatGPT 3.5 or higher drives the evolution of ML tools applications in ESG-related research. However, ML is dubbed a "black box" in accounting academic research because of its opaque nature and lack of transparency in decision-making. Specifically, traditional statistical models are easily interpretable, providing clear explanations of variable relationships, whereas complex ML algorithms pose challenges in understanding their inner workings and prediction logic. In addition, researchers should be serious and aware of concrete analysis of concrete problems when they consider applying ML in their research. For example, topic models (e.g. LDA) face difficulty in segregating ESG-related information into distinct topics due to the minuscule ESG information in annual reports compared to the vast volume of financial information they contain, but a word embedding model may be applicable.

The implementation of ML in ESG reporting and evaluation is currently in its initial phase. Nevertheless, it holds the promise of transforming the process of collecting, analyzing and utilizing ESG data. By harnessing ML, accounting researchers can uncover profound insights into the correlation between ESG practices and financial performance, leading to a more evidence-based comprehension of ESG's influence on business communities and accounting regulators. Yet, researchers must approach the application of ML in their research with seriousness and awareness, focusing on concrete analysis of specific problems.

4. Conclusions

ESG is a framework that incorporates environmental, social and governance factors into investment decisions, driving a significant increase in sustainable investment assets. While companies actively participate in ESG programs, challenges persist in ESG disclosure, including difficulties in quantifying unstructured data, lack of reporting guidelines and concerns over materiality and reliability. ESG ratings' credibility and consistency have been questioned, leading to increased interest in academic research to explore the determinants and consequences of ESG disclosures. ML tools have emerged as a means to enhance ESG quality assessment, focusing on disclosure, ratings and other ESG-related settings. In accounting research, the traditional logistic method aims to prioritize causal inference and minimize bias resulting from model misspecification, emphasizing high accuracy while enforcing predetermined relationships between variables to represent the underlying theory precisely.

Conversely, ML methods offer significant advantages over traditional approaches by efficiently handling large volumes of unstructured data commonly found in modern accounting datasets. ML's ability to capture complex patterns and non-linear relationships between variables, a challenge for traditional methods with rigid assumptions, contributes to its superiority. Furthermore, ML methods demonstrate superior prediction accuracy and classification performance, making them well-suited for tasks such as fraud detection, credit risk assessment and financial forecasting. The unique capability of ML models to automatically identify and utilize feature interactions allows a better understanding of variable interplay in the data. With greater adaptability and autonomy in discovering relevant features, ML methods are highly suitable for addressing intricate accounting problems with diverse and evolving datasets. In conclusion, adopting ML methods in accounting research brings numerous benefits, surpassing traditional logistic or simple linear regression approaches, as researchers can efficiently handle large, unstructured, non-numeric data, capture intricate relationships and achieve superior predictive performance, leading to more insightful and accurate analyses in the field of financial and nonfinancial reporting.

For future research, researchers could consider more applying word embedding models, topic models, generative AI and other ML models in ESG-related research due to corporate ESG information's non-numeric feature. Those ML models hold great promise in ESG research within the accounting field, enhancing the analysis of unstructured ESG data from various sources, providing multi-faceted insights into companies' ESG performance, facilitating materiality assessment and prioritizing significant ESG factors. They also could play a vital role in detecting greenwashing and ESG misreporting, promoting transparency and accountability through language cue identification in ESG disclosures. Additionally, those ML models can be used to conduct thematical ESG sentiment analysis in financial markets. This can help to understand investor behavior and assess risk. Nevertheless, researchers should be aware of the ML's "black box" feature, and carefully adopt the ML tools following the rule: specific analysis of specific issues.

Notes

1. https://www.unepfi.org/fileadmin/events/2004/stocks/who_cares_wins_global_compact_2004.pdf
2. https://www.msci.com/esg-and-climate-methodologies
3. Google incorporates Google BERT, an AI language model, into its search results to enhance their understanding. Despite its complexity, the purpose of Google BERT is straightforward: to improve Google's comprehension of the context within users. By utilizing NLP, natural language understanding and sentiment analysis, BERT employs AI techniques to analyze each word in a search query in relation to the entire sentence.
4. A transformer model is a type of neural network that acquires semantic context and understanding by observing relationships among sequential data elements, such as the words in a sentence. These models employ a dynamic set of mathematical techniques known as attention or self-attention to detect intricate connections, even between distant elements in a sequence, and determine how they influence and rely on each other.
5. DistilBERT, a lightweight variant of BERT developed by Hugging Face, offers comparable performance to BERT while requiring fewer parameters and shorter training times, making it an ideal choice for resource-limited environments like mobile and edge devices.
6. RoBERTa, similar to BERT, is a transformer-based language model that employs self-attention for processing input sequences and generating contextualized word representations. However, RoBERTa distinguishes itself by being trained on a significantly larger dataset and utilizing an improved training methodology compared to BERT.
7. The ESG team consists of 165 analysts, monitors approximately 1,700 European companies and has been generating ESG scores since 2002, offering both "ESG scores" and "ESG combined scores," along with corresponding literal ratings for each company, providing 178 ESG indicators.
8. The ESG team has approximately 185 dedicated analysts responsible for evaluating about 1,500 European companies, and their ESG rating time series spans over a 20-year period, providing 172 ESG variables.
9. Tree-based models belong to the family of supervised ML techniques that mainly handle classification and regression tasks by constructing a tree-like structure to determine the target variable's class or value based on the input features.
10. Word2Vec is a computationally efficient, two-layer neural network that learns word embeddings by mapping words from a large corpus to a vector space, where words with similar linguistic contexts in the corpus are located close to each other.
11. The financial API enables ESG analysts to access stock prices and market trends, providing real-time market data and trading functionality to improve their assessment of a company's ESG performance.
12. The data API encompassing diverse ESG data, like greenhouse gas emissions, labor practices and executive compensation, empowers ESG analysts to evaluate and rate companies' ESG performance.
13. The NLP API extracts unstructured content from news articles, social media posts and corporate reports, allowing ESG analysts to acquire extra insights on a company's ESG performance and merging this data into its scoring algorithm.
14. The data on shareholder proposals come from the Proxy Monitor database. https://www.proxymonitor.org/
15. Governance and Accountability Institute Inc., 2022 Sustainability Reporting in Focus is seen via https://www.ga-institute.com/research/ga-research-directory/sustainability-reporting-trends/2022-sustainability-reporting-in-focus.html
16. Ernst and Young "How can corporate reporting bridge the ESG trust gap?" is seen via https://assets.ey.com/content/dam/ey-sites/ey-com/en_gl/topics/assurance/assurance-pdfs/ey-global-reporting-survey-report-2022.pdf

17. Greenwashing is the act of deceptively or misleadingly conveying a false impression of environmental responsibility or sustainability by individuals, organizations, or companies. One classic example of greenwashing is the Volkswagen emissions scandal. Volkswagen confessed to manipulating emission test on by installing a "defeat" device in certain vehicles. The installed software could detect when the vehicle was undergoing an emissions test and adjust its performance to lower emissions.
18. See GHGRP via https://www.epa.gov/ghgreporting
19. Refer to the research by Rouen *et al.* (2023) about baseline words and phrases for GHG topics used to seed the ML algorithm and quantify the text for material topics (i.e. the SASB has a standard for GHG within that industry).
20. See the Climate and ESG Task Force Enforcements via https://www.sec.gov/securities-topics/enforcement-task-force-focused-climate-esg-issues
21. See H.R. 1,187 details via https://www.congress.gov/bill/117th-congress/house-bill/1187

References

Abeysekera, A.P. and Fernando, C.S. (2020), "Corporate social responsibility versus corporate shareholder responsibility: a family firm perspective", *Journal of Corporate Finance*, Vol. 61, 101370.

Aiba, Y., Ito, T. and Ibe, Y. (2019), "Network structure in ESG ratings suggests new corporate strategies: evolving AI technology to quantify qualitative data", *Securities Analysts Journal*.

Al-Hadi, A., Chatterjee, B., Yaftian, A., Taylor, G. and Hasan, M.M. (2017), "Corporate social responsibility performance, financial distress and firm life cycle: evidence from Australia", *Accounting and Finance*, Vol. 59 No. 2, pp. 961-989.

Allen, K. (2018), "Lies, damned lies and ESG rating methodologies", *Financial Times*, December 6, 2018.

Amel-Zadeh, A. and Serafeim, G. (2018), "Why and how investors use ESG information: evidence from a global survey", *Financial Analysts Journal*, Vol. 74 No. 3, pp. 87-103.

Baldini, M., Maso, L.D., Liberatore, G., Mazzi, F. and Terzani, S. (2018), "Role of country- and firm-level determinants in environmental, social, and governance disclosure", *Journal of Business Ethics*, Vol. 150, pp. 79-98.

Bao, Y., Ke, B., Li, B., Yu, Y.J. and Zhang, J. (2020), "Detecting accounting fraud in publicly traded US firms using a machine learning approach", *Journal of Accounting Research*, Vol. 58 No. 1, pp. 199-235.

Barboza, F., Kimura, H. and Altman, E. (2017), "Machine learning models and bankruptcy prediction", *Expert Systems with Applications*, Vol. 83, pp. 405-417.

Barnea, A. and Rubin, A. (2010), "Corporate social responsibility as a conflict between shareholders", *Journal of Business Ethics*, Vol. 97, pp. 71-86.

Baron, D.P. (2008), "Managerial contracting and corporate social responsibility", *Journal of Public Economics*, Vol. 92 No. 1, pp. 268-288.

Berg, F., Kölbel, J.F. and Rigobon, R. (2022), "Aggregate confusion: the divergence of ESG ratings", *Review of Finance*, Vol. 26 No. 6, pp. 1315-1344.

Bertomeu, J., Cheynel, E., Floyd, E. and Pan, W. (2021), "Using machine learning to detect misstatements", *Review of Accounting Studies*, Vol. 26, pp. 468-519.

Bingler, J.A., Kraus, M., Leippold, M. and Webersinke, N. (2022), "Cheap talk and cherry-picking: what ClimateBert has to say on corporate climate risk disclosures", *Finance Research Letter*, Vol. 47, 102776.

Booker, A., Curtis, A. and Richardson, V.J. (2023), "Investor disagreement, disclosure processing costs, and trading volume evidence from social media", *The Accounting Review*, Vol. 98 No. 1, pp. 109-137.

Borghesi, R., Houston, J.F. and Naranjo, A. (2014), "Corporate socially responsible investments: CEO altruism, reputation, and shareholder interests", *Journal of Corporate Finance*, Vol. 26, pp. 164-181.

Briscoe-Tran, H. (2023), "Do employees have useful information about firms' esg practices?", Working paper, University of Alberta.

Brown, W., Helland, E. and Smith, J. (2006), "Is corporate philanthropy an agency cost?", *Journal of Corporate Finance*, Vol. 12, pp. 855-877.

Bui, B., Houqe, M.N. and Zaman, M. (2020), "Climate governance effects on carbon disclosure and performance", *British Accounting Review*, Vol. 52, 100880.

Burke, J.J., Hoitash, R. and Hoitash, U. (2019), "Auditor response to negative media coverage of client environmental, social, and governance practices", *Accounting Horizons*, Vol. 33 No. 1, pp. 1-23.

Byard, D., Li, Y. and Yu, Y. (2011), "The effect of mandatory IFRS adoption on financial analysts' information environment", *Journal of Accounting Research*, Vol. 49 No. 1, pp. 69-96.

Chae, B. and Park, E. (2018), "Corporate social responsibility (CSR): a survey of topics and trends using Twitter data and topic modeling", *Sustainability*, Vol. 10 No. 7, pp. 1-20.

Chatterji, A. and Toffel, M. (2010), "How firms respond to being rated", *Strategic Management Journal*, Vol. 31 No. 9, pp. 917-945.

Chatterji, A., Durand, R., Levine, D. and Touboul, S. (2016), "Do ratings of firms converge? Implications for managers, investors and strategy researchers", *Strategic Management Journal*, Vol. 37 No. 8, pp. 1597-1614.

Cheng, B., Ioannou, I. and Serafeim, G. (2014), "Corporate social responsibility and access to finance", *Strategic Management Journal*, Vol. 35 No. 1, pp. 1-23.

Christensen, D.M., Serafeim, G. and Sikochi, A. (2022), "Why is corporate virtue in the eye of the beholder? The case of ESG ratings", *The Accounting Review*, Vol. 97 No. 1, pp. 147-175.

Clarkson, P., Li, Y., Pinnuck, M. and Richardson, G. (2015), "The valuation relevance of greenhouse gas emissions under the European Union carbon emissions trading scheme", *European Accounting Review*, Vol. 24 No. 3, pp. 551-580.

Cowen, S.S., Ferrer, L.B. and Parker, L.D. (1987), "The impact of corporate characteristics on social responsibility disclosure: a typology and frequency-based analysis", *Accounting, Organizations and Society*, Vol. 12 No. 2, pp. 111-122.

Dai, X., Gao, F., Lisic, L.L. and Zhang, I.X. (2023), "Corporate social performance and the managerial labor market", *Review of Accounting Studies*, Vol. 28, pp. 307-339.

Dang, H.N., Hoang, K., Vu, V.T. and Nguyen, L.V. (2021), "Do socially responsible firms always disclose high-quality earnings? Evidence from an emerging socialist economy", *Asian Review of Accounting*, Vol. 29 No. 3, pp. 291-306.

Davis, A.K., Guenther, D.A., Krull, L.K. and Williams, B.M. (2016), "Do socially responsible firms pay more taxes?", *The Accounting Review*, Vol. 91 No. 1, pp. 47-68.

Del Vitto, A., Marazzina, A. and Stocco, D. (2023), "ESG ratings explainability through machine learning techniques", *Annals of Operations Research*, (forthcoming).

Deloitte Insights (2022), "Ingraining sustainability in the next era of ESG investing", available at: https://www2.deloitte.com/us/en/insights/industry/financial-services/esg-investing-andsustainability.html

Dhaliwal, D.S., Radhakrishnan, S., Tsang, A. and Yang, Y.G. (2012), "Nonfinancial disclosure and analyst forecast accuracy: international evidence on corporate social responsibility disclosure", *The Accounting Review*, Vol. 87 No. 3, pp. 723-759.

Dube, S. and Zhu, C. (2021), "The disciplinary effect of social media: evidence from firms' responses to Glassdoor reviews", *Journal of Accounting Research*, Vol. 59 No. 5, pp. 1783-1825.

Edmans, A. (2023), "The end of ESG", *Financial Management*, Vol. 52 No. 1, pp. 3-17.

Erel, I., Stern, L.H., Tan, C. and Weisbach, M.S. (2021), "Selecting directors using machine learning", *Review of Financial Studies*, Vol. 34 No. 7, pp. 3226-3264.

Föhr, T.L., Schreyer, M., Juppe, T.A. and Marten, K. (2023), "Assuring sustainable futures: auditing sustainability reports using ai foundation models", Working paper, Ulm University and University of St. Gallen.

Friederich, D., Kaack, L.H., Luccioni, A. and Steffen, B. (2021), "Automated identification of climate risk disclosures in annual corporate reports", *[Paper presentation]. ICML*, July 23, 2021, Online available.

Gao, L. and Zhang, J.H. (2015), "Firms' earnings smoothing, corporate social responsibility, and valuation", *Journal of Corporate Finance*, Vol. 32, pp. 108-127.

Gibson Brandon, R., Krueger, P. and Schmidt, P.S. (2021), "ESG rating disagreement and stock returns", *Financial Analysts Journal*, Vol. 77 No. 4, pp. 104-127.

Global Sustainable Investment Alliance (2020), "Global sustainable investment review 2020".

Goloshchapova, I., Poon, S., Pritchard, M. and Reed, P. (2019), "Corporate social responsibility reports: topic analysis and big data approach", *European Journal of Finance*, Vol. 25 No. 17, pp. 1637-1654.

Grewal, J., Riedl, E.J. and Serafeim, G. (2019), "Market reaction to mandatory nonfinancial disclosure", *Management Science*, Vol. 65 No. 7, pp. 3061-3084.

Griffin, P.A., Lont, D.H. and Sun, Y. (2017), "The relevance to investors of greenhouse gas emission disclosures", *Contemporary Accounting Research*, Vol. 34 No. 2, pp. 1265-1297.

Guenther, D.A., Peterson, K., Searcy, J. and Williams, B.M. (2023), "How useful are tax disclosures in predicting effective tax rates? A machine learning approach", *The Accounting Review*, Vol. 98 No. 5, pp. 1-26.

Guo, J. and Yu, Y. (2022), "CSR committees, politicians and CSR efforts", *Asian Review of Accounting*, Vol. 30 No. 3, pp. 297-313.

Hegde, S.P. and Mishra, D.R. (2019), "Married CEOs and corporate social responsibility", *Journal of Corporate Finance*, Vol. 58, pp. 226-246.

Henry, E. and Leone, A.J. (2016), "Measuring qualitative information in capital markets research: comparison of alternative methodologies to measure disclosure tone", *The Accounting Review*, Vol. 91 No. 1, pp. 153-178.

Hoi, C.K., Wu, Q. and Zhang, H. (2013), "Is corporate social responsibility (CSR) associated with tax avoidance? Evidence from irresponsible CSR activities", *The Accounting Review*, Vol. 88 No. 6, pp. 2025-2059.

Huang, A.H., Zang, A.Y. and Zheng, R. (2014), "Evidence on the information content of text in analyst reports", *The Accounting Review*, Vol. 89 No. 6, pp. 2151-2180.

Huang, A.H., Wang, H. and Yang, Y. (2023), "FinBERT: a large language model for extracting information from financial text", *Contemporary Accounting Research*, Vol. 40 No. 2, pp. 806-841.

Hunt, J.O., Rosser, D.M. and Rowe, S.P. (2021), "Using machine learning to predict auditor switches: how the likelihood of switching affects audit quality among non-switching clients", *Journal of Accounting and Public Policy*, Vol. 40 No. 5, 106785.

Iliev, P. and Roth, L. (2023), "Director expertise and corporate sustainability", *Review of Finance*, (forthcoming).

Ioannou, I., Li, S.X. and Serafeim, G. (2016), "The effect of target difficulty on target completion: the case of reducing carbon emissions", *The Accounting Review*, Vol. 91 No. 5, pp. 1467-1492.

Jain, Y., Gupta, S., Yalciner, S., Joglekar, Y., Khetan, P. and Zhang, T. (2023), "Overcoming complexity in ESG investing: The role of generative AI integration in identifying contextual ESG factors", Working paper, University of Illinois at Urbana-Champaign.

Jiang, L., Gu, Y. and Dai, J. (2023), "Environmental, social, and governance taxonomy simplification: a hybrid text mining approach", *Journal of Emerging Technologies in Accounting*, Vol. 20 No. 1, pp. 305-325.

Khlif, H., Ahmed, K. and Souissi, M. (2017), "Ownership structure and voluntary disclosure: a synthesis of empirical studies", *Australian Journal of Management*, Vol. 42 No. 3, pp. 376-403.

Kiesel, F. and Lücke, F. (2019), "ESG in credit ratings and the impact on financial markets", *Financial Markets, Institutions and Instruments*, Vol. 28 No. 3, pp. 263-290.

Kim, Y., Park, M.S. and Wier, B. (2012), "Is earnings quality associated with corporate social responsibility?", *The Accounting Review*, Vol. 87 No. 3, pp. 761-796.

Kimbrough, M.D., Wang, X., Wei, S. and Zhang, J. (2022), "Does voluntary ESG reporting resolve disagreement among ESG rating agencies?", *European Accounting Review*, (forthcoming).

King, G., Lam, P. and Roberts, M.E. (2017), "Computer-assisted keyword and document set discovery from unstructured text", *American Journal of Political Science*, Vol. 61 No. 4, pp. 971-988.

Klusak, P., Agarwala, M., Burke, M., Kraemer, M. and Mohaddes, K. (2023), "Rising temperatures, falling ratings: the effect of climate change on sovereign creditworthiness", *Management Science*, (forthcoming).

KPMG (2022), "Survey of sustainability reporting 2022", available at: https://assets.kpmg.com/content/dam/kpmg/se/pdf/komm/2022/Global-Survey-of-Sustainability-Reporting-2022.pdf

Lalova, K. (2023), "The value of employee morale in mergers and acquisitions: Evidence from Glassdoor", Working paper, Michigan State University.

Lanza, A., Bernardini, E. and Faiella, I. (2020), "Mind the gap! Machine learning, ESG metrics and sustainable investment", Bank of Italy Occasional Paper, No.561.

Li, F. (2010), "The information content of forward-looking statements in corporate filings – a Naïve Bayesian machine learning approach", *Journal of Accounting Research*, Vol. 48 No. 5, pp. 1049-1102.

Li, K., Mai, F., Shen, R. and Yan, X. (2021), "Measuring corporate culture using machine learning", *Review of Financial Studies*, Vol. 34 No. 7, pp. 3265-3315.

Lin, Y., Shen, R., Wang, J. and Yu, Y.J. (2023), "Global evolution of environmental and social disclosure in annual reports", Working paper, University of Virginia.

Liu, W.P., Yen, M.F. and Wu, T.Y. (2022), "Report users' perceived sentiments of key audit matters and firm performance: evidence from a deep learning-based natural language processing approach", *Journal of Information Systems*, Vol. 36 No. 3, pp. 191-209.

Lougee, B. and Wallace, J. (2008), "The corporate social responsibility (CSR) trend", *Journal of Applied Corporate Finance*, Vol. 20, pp. 96-108.

Loughran, T. and McDonald, B. (2016), "Textual analysis in accounting and finance: a survey", *Journal of Accounting Research*, Vol. 54 No. 4, pp. 1187-1230.

Luccioni, A., Baylor, E. and Duchene, N. (2020), "Analyzing sustainability reports using natural language processing", *arXivLabs*, arXiv:2011.08073.

Martin, P.R. and Moser, D.V. (2016), "Managers' green investment disclosures and investors' reaction", *Journal of Accounting and Economics*, Vol. 61 No. 1, pp. 239-254.

McGuinness, P.B., Vieito, J.P. and Wang, M. (2017), "The role of board gender and foreign ownership in the CSR performance of Chinese listed firms", *Journal of Corporate Finance*, Vol. 42, pp. 75-99.

Michelon, G., Rodrigue, M. and Trevisan, E. (2020), "The marketization of a social movement: activists, shareholders and CSR disclosure", *Accounting, Organizations and Society*, Vol. 80, 101074.

Naughton, J.P., Wang, C. and Yeung, I. (2019), "Investor sentiment for corporate social performance", *The Accounting Review*, Vol. 94 No. 3, pp. 401-420.

Perols, J. (2011), "Financial statement fraud detection: an analysis of statistical and machine learning algorithms", *Auditing: A Journal of Practice and Theory*, Vol. 30 No. 2, pp. 19-50.

Pinnuck, M., Ranasinghe, A., Soderstrom, N. and Zhou, J. (2021), "Restatement of CSR reports: frequency, magnitude, and determinants", *Contemporary Accounting Research*, Vol. 38 No. 3, pp. 2376-2416.

Porter, M. and Kramer, M. (2006), "Strategy and society: the link between competitive advantage and corporate social responsibility", *Harvard Business Review*, Vol. 84 No. 12, pp. 78-92.

Rastogi, S., Singh, K. and Kanoujiya, J. (2023), "Firm's value and ESG: the moderating role of ownership concentration and corporate disclosures", *Asian Review of Accounting*, (forthcoming).

Rouen, E., Sachdeva, K. and Yoon, A. (2023), "The evolution of ESG reports and the role of voluntary standards", Working paper, Rice University, and Northwestern University, Harvard Business School.

Ryou, J.W., Tsang, A. and Wang, K.T. (2022), "Product market competition and voluntary corporate social responsibility disclosures", *Contemporary Accounting Research*, Vol. 39 No. 2, pp. 1215-1259.

Sautner, Z., Lent, L.V., Vilkov, G. and Zhang, R. (2023), "Firm-level climate change exposure", *The Journal of Finance*, Vol. 78 No. 3, pp. 1449-1498.

Serafeim, G. and Yoon, A. (2023), "Stock price reactions to ESG news: the role of ESG ratings and disagreement", *Review of Accounting Studies*, Vol. 28, pp. 1500-1530.

SustainAbility (2020), "Rate the raters 2020: investor survey and interview results", available at: https://www.sustainability.com/globalassets/sustainability.com/thinking/pdfs/sustainability-ratetheraters2020-report.pdf

Task Force on Climate-related Financial Disclosures (2022), "2022 TCFD status report", available at: https://assets.bbhub.io/company/sites/60/2022/10/2022-TCFD-Status-Report.pdf

Wanderley, L.S.O., Lucian, R., Farache, F. and Filho, J. M.de S. (2008), "CSR information disclosure on the Web: a context-based approach analysing the influence of country of origin and industry sector", *Journal of Business Ethics*, Vol. 82, pp. 369-378.

Welch, K. and Yoon, A. (2023), "Do high-ability managers choose ESG projects that create shareholder value? Evidence from employee opinions", *Review of Accounting Studies*, (forthcoming).

Zhao, D. and Huang, X. (2021), "Environmental, social and governance factors and assessing firm value: valuation, signaling and stakeholder perspectives", *Accounting and Finance*, Vol. 62 No. 1, pp. 1983-2010.

Appendix

Literature of ESG

Global Sustainable Investment Alliance (2020) reports that sustainable investment remains prominent in the global investment industry, with assets under management reaching US$ 35.3 trillion, representing a 15% growth over two years (2018–2020). Notably, sustainable investment assets in the United States have experienced a significant rise, with a remarkable 42% increase from 2018 to 2020. A recent Deloitte Insights study (2022) shows that ESG-mandated assets, which encompass professionally managed assets that incorporate ESG considerations in investment selection or file shareholder resolutions on ESG issues in publicly traded companies, are expected to comprise approximately 50% of all professionally managed assets worldwide by 2024.

Meanwhile, investors are placing a growing emphasis on integrating ESG within companies' operational strategies, as evident from the content of their proposals [14]. Responding to investors' ESG demands, firms are witnessed to be engaged actively in ESG practice. Currently, Governance & Accountability Institute Inc., an ESG consultancy, reports that 96% of companies in the S&P 500 and 81% of companies in the Russell 1,000 Index published sustainability reports in 2021 [15]. Likewise, according to a survey released by KPMG (2022), 96% of the top 250 global companies, based on revenue, publicly disclose information on ESG or sustainability topics. Over 50% of the investors surveyed regard firms' ESG disclosures as a bonanza for assessing their ESG performance based on a report conducted by SustainAbility (2020).

However, Ernst and Young's 2022 survey indicates that around 80% of investors express their concerns that numerous companies struggle to effectively communicate the reasons behind long-term sustainability investments, making it challenging for them to assess the viability and potential of such investments [16]. The existing challenges in firms' ESG disclosure exacerbate public concerns. Specifically, most of the information in ESG reports is challenging to calculate or quantify (i.e. unstructured data), which contrasts with financial reports that primarily revolve around monetary values. For example, insurance companies use images to estimate the damage caused by a tornado. Some firms use social media to evaluate the degree of union support within their workforce.

The lack of reporting guidelines and comparability is another issue. A multitude of ESG frameworks, such as the Task Force on Climate-related Financial Disclosures (TCFD), Sustainability Accounting Standards Board (SASB), Global Reporting Initiative (GRI), Carbon Disclosure Project (CDP), and the EU Taxonomies, have their own standards, inhibiting the comparability of ESG information for investors. The absence of materiality is exemplified by the consideration that water consumption holds substantial significance in assessing a fish processing facility. Conversely, this metric may possess limited pertinence when prospective investors are deliberating investments in financial institutions, compounded by a deficiency in reliability. Take, for example, greenwashing represents an even more extreme form of misrepresenting ESG efforts [17].

To address public concerns regarding ESG disclosure, numerous nations have implemented regulations that demand public companies to disclose their ESG performance. For instance, the Greenhouse Gas Reporting Program (GHGRP), administrated by the US Environmental Protection Agency, was designed with the primary purpose of guiding potential future greenhouse gas (GHG) policies, mandatorily requiring thousands of facilities in the US to report their annual emissions and production activity related to GHGs [18, 19]. Besides, in March 2021, the U. S. Securities and Exchange Commission launched the Climate and ESG Task Force to detect potential breaches of current disclosure regulations pertaining to climate risks and ESG strategies [20]. Concurrently, the US House passed the Corporate Governance Improvement and Investor Protection Act (H.R. 1,187), compelling SEC to promulgate regulations that establish clear definitions for ESG metrics [21].

The financial market and regulators increasingly focus on the ESG issue, spurring the assessment of firms' ESG behaviors. ESG rating agencies arise at the right moment, meticulously examining businesses and evaluating their corporate ESG performance with their own different metrics. Additionally, we have witnessed the emergence and evolution of ESG rating agencies, such as Morgan Stanley Capital International (MSCI), Institutional Shareholder Services (ISS ESG), Sustainalytics, Bloomberg, Refinitiv (formerly known as Thomson Reuters), etc. Similar to credit rating agencies, these ESG rating agencies act as the main intermediaries of ESG information in financial markets, influencing the stakeholders' decision-making. However, skeptics express concern that ESG ratings may not effectively differentiate socially responsible firms and argue that these ratings can generate metrics that are frequently invalid and potentially misleading to stakeholders.

Anecdotal evidence agrees with those concerns. For instance, Allen (2018) expresses her doubts about the inconsistency of ESG ratings among rating agencies by "Investors need to be clear about what the methodology they choose is actually measuring, and why. Otherwise, ESG scoring risks creating a false sense of confidence among investors who don't really understand what lies behind the numbers– and therefore don't really understand what they're buying" and takes Tesla as an example that Tesla received a low rating for global auto ESG from FTSE, whereas MSCI rated it as the best. In Sustainalytics rankings, Tesla's position was closer to the middle of the pack. Interestingly, one researcher argues that diverse ratings on firms' ESG performance provide richer insights, as ESG-by-numbers investors seek a single ESG rating, while mainstream investors prefer thorough analysis and evaluation from various sources before making portfolio decisions (Edmans, 2023).

A plethora of theories seek to answer the question of what factors affect firms' ESG disclosure practices. Stakeholder theory suggests that managers are responsible for striking a balance among the competing interests of various stakeholders. Building on this theory, researchers (such as Kim *et al.*, 2012; Ioannou *et al.*, 2016; Martin and Moser, 2016; Naughton *et al.*, 2019) integrate the concept of ESG with shareholder wealth maximization. Firms are expected to have good sustainable performance to meet the ESG demands from stakeholders, such as suppliers, investors, employees, and others. Agency theory (Christensen *et al.*, 2022) explains that the self-interest of managers may engage in biased ESG disclosures to create an impression of being better performers in CSR than they truly are. This manipulation allows them to enhance the reputation of their firms, potentially benefiting their personal and professional interests.

According to the legitimacy theory, it can be anticipated that companies with poorer ESG performance would tend to provide more extensive off-setting or positive ESG disclosures in their financial reports, mitigating the negative perception of their ESG performance and showcasing their efforts to address ESG concerns (Michelon *et al.*, 2020; Pinnuck *et al.*, 2021). Lastly, Ryou *et al.* (2022) link product market competition and ESG disclosures with the explanation of the signaling theory that firms with ESG disclosures are exposed to competitive advantages over their competitors, signaling good product quality.

Corresponding author
Joseph H. Zhang can be contacted at: jzhang5@memphis.edu