What kind of ESG is profitable? Connecting company performance to ESG terms in financial reports

Luka Andrenšek
trovato@corporation.com
Jožef Stefan Institute
Ljubljana, Slovenia

Katarina Sitar Šuštar
University of Ljubljana
Ljubljana, Slovenia
katarina.sitar@ef.uni-lj.si

Senja Pollak
Jožef Stefan Institute
Ljubljana, Slovenia
senja.pollak@ijs.si

Matthew Purver
Jožef Stefan Institute
Ljubljana, Slovenia
matthew.purver@ijs.si

ABSTRACT

In this paper, we examine the relationship between the discussion of Environmental, Social and Governance (ESG) in companies' annual financial reports and their financial performance. Specifically, we analyse the companies' use of specific ESG terms alongside the performance metric, sector-normalized Return on Assets (ROA). Our motivation is to determine whether companies frequently mentioning terms such as "gender", "equality", "talent", and "innovation" in their reports demonstrate a higher annual ROA compared to those that rarely used these terms. To explore this, we used existing datasets with reports and performance metrics from 348 companies, covering the years from 2009 to 2021. In order to better examine differences, we then selected companies whose ROA significantly differed from the average (either higher or lower), allowing for a more pronounced examination of the impact of ESG term usage on financial performance. The filtered dataset consisted of 107 companies, with a total of 427 reports; split into two sections representing higher and lower performing companies. We then used an existing list of ESG terms derived from a range of separate data sources, and applied a basic statistical n-gram language model to extract the probabilities of each ESG term's occurrence in each of the higher- and lower-performing dataset sections. Results show that while certain sets of ESG concepts correlate with higher financial performance, others do the opposite, and give some initial interpretation into the light this sheds on company reporting behaviour.

KEYWORDS

financial report analysis, language modelling, environmental, social and governance reporting

1 INTRODUCTION & RELATED WORK

There is increasing interest in the behaviour of companies in the area of Environmental, Social and Governance (ESG) criteria, including a company's environmental impact (Environmental), relationships with the community including employees, suppliers and customers (Social), and leadership structures including executive pay and shareholder rights (Governance). Although until recently, ESG analyses were almost entirely performed manually by experts (see e.g. [10]), there has been a large amount of work in the last few years on applying computational machine learning and statistical methods to ESG analysis (see e.g. the recent review by Lim [9]).

However, much of this analysis examines numerical company performance data and categorical metadata; our interest is in developing and applying natural language processing (NLP) technologies to not only help automate analyses, but allow understanding of how human actors discuss and understand the important and meaning of ESG aspects.

Application of NLP in finance is not new: for example, topic modelling has been used to predict company performance and investigate strategies [14, 7]. Recent work also includes application to ESG aspects: Nugent et al. [12] automatically extract news about ESG controversies, and Lee et al. [8] analyse sentiment on ESG issues. Closer to our interests, Purver et al. [13] investigated how the use of ESG terms by companies has changed over time. By analysing and annotating a set of existing resources, they defined a set of 93 ESG terms categorised into 5 core ESG areas. They then showed how these terms can be used to analyse changes in reporting, by analysing a collection of company annual reports, collated over a period of 8 years, using language modelling and distributional methods to reveal changes in the frequency and in the usage of the ESG terms.

Here, we are interested not in changes in ESG discussion over time, but in whether and how the reporting of ESG aspects is connected to financial performance. We take Purver et al. [13]'s resources and methods as a starting point, but augment the financial report text data with available metadata on financial performance, allowing us to compare how ESG reporting varies between more and less well-performing companies.

2 DATA AND METHODS

2.1 Hypotheses

In general, we expect increased probability of appearance of ESG terms in the annual reports from the more profitable firms, based on a number of factors. In general, overall high ESG performing companies exhibit high financial performance [1, 5]; although we note that the link between high ESG score performance and mention of ESG terms is not guaranteed to be straightforward. More specifically, during the period between 2010-2020 analysed here, there was a growing emphasis on corporate social responsibility (CSR) and sustainability. Investors, consumers, and other stakeholders increasingly prioritised companies that demonstrated a commitment to innovation, diversity, and environmental sustainability [11, 2]. Busru and Shanmugasundaram [3] find that firms closely engaging in fostering innovation, attracting top talent,

Information Society 2024, 7–11 October 2024, Ljubljana, Slovenia

https://doi.org/10.70314/is.2024.sikdd.3

Year	# Reports	# Words
2012	178	12.5M
2013	181	14.0M
2014	184	15.0M
2015	196	16.3M
2016	198	17.5M
2017	200	18.4M
2018	200	19.6M
2019	202	21.2M
total	1539	134.6M

Table 1: Number of annual reports available by year

promoting gender and diversity initiatives, could confer a competitive advantage over the industry peers. Furthermore, some policy and regulatory changes (e.g. the 2018 UK Corporate Governance Code, the 2014 EU Directive on Non-Financial Reporting, Carbon Disclosure Project (CDP)) directly or indirectly encouraged companies to address issues related to diversity, gender equality, and environmental sustainability.

2.2 Data and pre-processing

To test this hypothesis, we build on the resources and methods of Purver et al. [13], who provide a dataset of annual reports from FTSE350 companies over the years 2012-2019, based on the FTSE350 list as of 25th April 2020 and obtained from the publicly accessible collection at www.annualreports.com. The reports are already converted to plain text, and we use their publicly available tools to tokenize the collection into words and build ngrams of length 1-4 padded with sentence start and end symbols; the dataset size is reported in Table 1 below (taken from [13]). We use their set of ESG terms, defined via a process of extracting candidate terms from a set of public ESG definitions and taxonomies, asking financial expert annotators to label them as to their representativeness as ESG terms and their ESG subcategory, and keeping the terms with high inter-annotator agreement (see [13] for details).

2.3 Financial performance analysis

The reports were then linked to financial indicators for the respective year and company. The data on company fundamentals was obtained from the Refinitiv EIKON Datastream.[1] Each entry contained annual financial indicators, as well as the companies' industry and sector codes. The main variable of interest was normalized, averaged *return on assets (ROA)* as defined below:[2]

$$\frac{\begin{array}{l} NetIncome - BottomLine \\ \quad +((InterestExpenseOnDebt - InterestCapitalized) \\ \qquad \times(1 - TaxRate)) \end{array}}{AverageOfLastYear'sAndCurrentYear'sTotalAssets}$$

After extracting financial reports with available ROA data, we categorized the financial reports into two groups, in order to examine differences in the associated reports' use of ESG terms. The distribution of ROA shows a heavy concentration around the mean, so in order to derive two distinctive groups we took the two extremes and excluded the central group around the mean. The 'negative' group comprised reports with a yearly ROA less than -0.2, indicating very poor performance. Conversely, the 'positive' group included reports with an ROA of at least 0.2, reflecting very good yearly performance.

Subsequently, we employed a statistical n-gram language model (using NLTK[3]) to analyze the occurrence of each ESG term. For each term, we calculated the probability of its occurrence in positive reports (p_+) and in negative reports (p_-), and the difference ($p_+ - p_-$). Terms with a large difference in these probabilities are more strongly associated with positive reports than with negative ones, and vice versa: terms with a large negative difference are common in negative reports but rare in positive ones. We conducted this analysis for both unigrams and bigrams.

3 RESULTS AND DISCUSSION

The results for 1- and 2-grams are shown in Figures 1 and 2 below (3- and 4-grams showed no clear interpretable associations).[4] As hypothesized, many ESG terms show a strong association with positive performance, with many of these being core terms associated with human resources (*innovation*, *talent*), with social aspects (*gender*, *diversity*), environmental aspects (*renewable*, *carbon footprint*, *environmental impact*) and overall ESG descriptors (*ethical*). However, many terms are conversely (and contrary to our general hypothesis) associated with negative performance, including, again, terms across various ESG categories including environmental (*carbon emissions*, *energy efficiency*, *greehouse*), human resources (*mental health*, *wellbeing*) and general ESG descriptors (*governance*).

However, by combining these terms with recent work in clustering and describing ESG terms [4], we can shed more light on which categories seem to be more positive and which more negative. Ferjancic et al. [4], using the same dataset and ESG term list [13], perform a further topic analysis using BERTopic [6], in which they derive 30 ESG-related topics and 6 higher-level clusters of ESG concepts; they then examine the correlations between these ESG topics and company ESG scores as obtained from external analysts. We align our ESG terms with Ferjancic et al. [4]'s 30 topics by matching against the words most associated with each topic (if a term appears in the top 10 words associated with a topic, we take the term and topic as aligned); we can then compare our positive/negative associations with Ferjancic et al. [4]'s correlations with company ESG scores. Table 2 shows this alignment for our most positive and negative bigram terms here, with the topic labels and an indication of the strength and direction of correlation with overall company ESG scores, as given by [4].

Given this, we see some systematic groupings. *Climate change*, as part of the 'climate risk and policy' topic, as well as *supply chain* and *human trafficking* as part of the 'human rights' topic, represent the themes that appear to be, across different industries, related to high company ESG scores. A similar observation holds for *gender balance*, *gender pay* and *environmental impact*, which all fall in a group of topics which are strongly and significantly correlated with high ESG scores throughout different industries. Overall high ESG performing companies exhibit high financial performance [1, 5], therefore our results for terms such as *climate change*, *supply chain* and *human trafficking* are not surprising: as indicators of topics associated with high ESG, they are good terms for tracking these ESG aspects associated with high financial performance.

[1] https://www.refinitiv.com
[2] We use this normalization and averaging to smooth and remove one-off effects.
[3] https://www.nltk.org/
[4] Note that these figures show differences in absolute probabilities: magnitudes are comparable within 1-grams, and within 2-grams, but not between 1- and 2-grams.

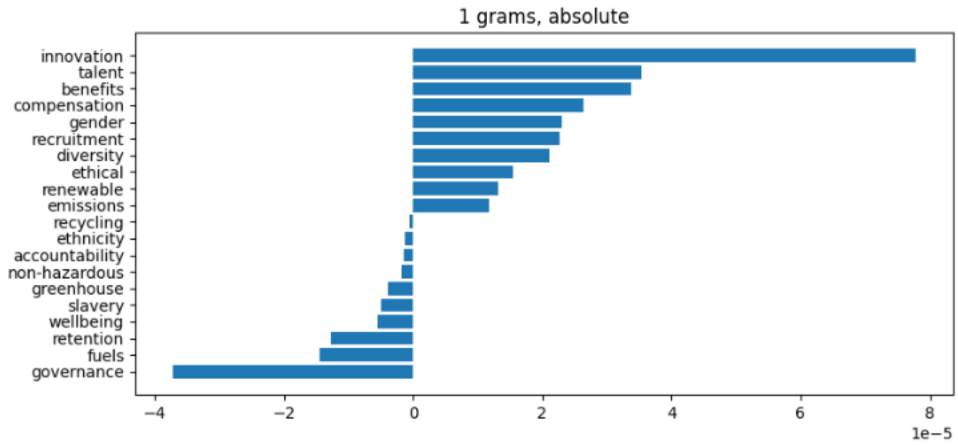

Figure 1: Difference in probability between positive and negative reports $p_+ - p_-$ for the most positive and negative unigram ESG terms.

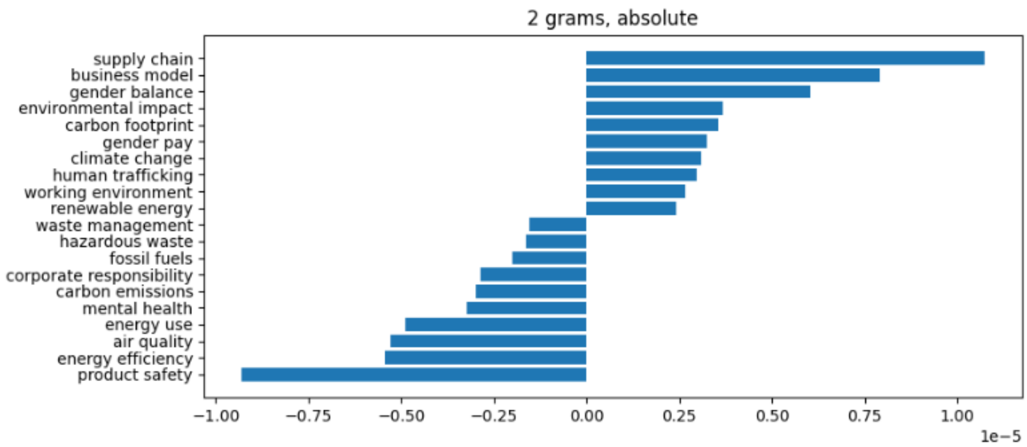

Figure 2: Difference in probability between positive and negative reports $p_+ - p_-$ for the most positive and negative bigram ESG terms.

Looking at the terms with low values which are associated with low RoA, *waste management* and *corporate responsibility* are associated with topics, for which in some industries proportion of these correlate with ESG scores significantly positively and in other industries this correlation is significantly negative. Based on overall correlation between ESG scores and topic proportions across different industries, these two topics are among the third of the topics for which negative correlation between the topic proportion and ESG score prevails. Due to the aforementioned correlation between ESG and financial performance it is therefore understandable that these terms are associated with mention in annual reports of companies with low RoA. Overly extensive discussion on specific topics (such as 'waste management' and 'corporate responsibility') can negatively impact ESG score (see [4]) which can by analogy of ESG and financial performance [1, 5] hold for companies with low RoA.

There is a surprising number of bigrams in both the high RoA and low RoA groups which seem to be associated with the same topic, namely 'climate footprint and energy management'. For companies with high RoA, these terms are *carbon footprint* and *renewable energy*, and for companies with low RoA, the terms are *fossil fuels*, *carbon emissions*, *energy use*, *air quality* and *energy efficiency*. It seems that better performing companies use *carbon footprint* instead of *carbon emissions*, and discuss more on the use of *renewable energy* than on *energy use*, *energy efficiency* and/or *fossil fuels*. In future work, we plan to analyse the use of these terms in more depth, including analysis of the lexical and topical contexts in which they appear, and adding techniques such as sentiment and topic analysis to shed more light on these distinctions.

ACKNOWLEDGEMENTS

The authors thank the reviewers for helpful suggestions, and acknowledge financial support from the Slovenian Research Agency for research core funding (No. P2-0103), as well as for funding of the research project *Quantitative and qualitative analysis of the unregulated corporate financial reporting* (No. J5-2554).

REFERENCES

[1] Nisar Ahmad, Asma Mobarek, and Naheed Nawazesh Roni. 2021. Revisiting the impact of ESG on financial performance of FTSE350 UK firms: static and dynamic panel data analysis. *Accounting, Corporate Governance & Business Ethics*. DOI: 10.1080/23311975.2021.1900500.

2 grams	Term/ROA correlation	Topic	Topic/ESG score correlation
Supply chain	+	Human rights	++
Business model	+	Customer services, People and culture	+; -
Gender balance	+	Diversity and inclusion	++
Environmental impact	+	General ESG	+
Carbon footprint	+	Climate footprint and energy management	=
Gender pay	+	Diversity and inclusion	++
Climate change	+	Climate risk and policy	++
Human trafficking	+	None directly related, in broader context in Human rights	++
Working environment	+	People and culture	-
Renewable energy	+	Climate footprint and energy management	=
Waste management	-	Waste management	–
Fossil fuels	-	No explicit match; contextually appears in Climate footprint and energy management	=
Corporate responsibility	-	Corporate governance	–
Carbon emissions	-	Climate footprint and energy management	=
Mental health	-	Health and safety	+
Energy use	-	Climate footprint and energy management	-
Air quality	-	No explicit match; contextually appears in Climate footprint and energy management	-
Energy efficiency	-	Climate footprint and energy management	-
Product safety	-	Health and safety	=

Table 2: Selected ESG terms with their ROA correlation direction (+/−), topic according to [4], and topic/ESG score correlation strength (+ + /+/= /−/−−) as calculated by [4].

[2] A. C. Amason and H. J. Sapienza. 2012. The effects of top management team size and interaction norms on cognitive and affective conflict. *Journal of Management*, 23, 495–516.

[3] S. A. Busru and G. Shanmugasundaram. 2017. Effects of innovation investment on profitability and moderating role of corporate governance: empirical study of indian listed firms. *Indian Journal of Corporate Governance*, 10, 97–117, 2. https://doi.org/10.1177/0974686217730938.

[4] Ursa Ferjancic et al. forthcoming. Textual analysis of corporate sustainability reporting and corporate ESG scores. under review. (Forthcoming).

[5] Gunnar Friede, Timo Busch, and Alexander Bassen. 2015. ESG and financial performance: aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment*, 5, 4, 210–233. DOI: 10.1080/20430795.2015.1118917.

[6] Maarten Grootendorst. 2022. BERTopic: neural topic modeling with a class-based TF-IDF procedure. (2022). https://arxiv.org/abs/2203.05794 arXiv: 2203.05794 [cs.CL].

[7] M. Jagannathan, D. Roy, and V. S. K. Delhi. 2022. Application of NLP-based topic modeling to analyse unstructured text data in annual reports of construction contracting companies. *CSI Transactions on ICT*, 10, 2, 97–106.

[8] H. Lee, S. H. Lee, K. R. Lee, and J. H. Kim. 2023. Esg discourse analysis through bertopic: comparing news articles and academic papers. *Computers, Materials & Continua*, 75, 3, 6023–6037.

[9] Tristan Lim. 2024. Environmental, social, and governance (esg) and artificial intelligence in finance: state-of-the-art and research takeaways. *Artificial Intelligence Review*, 57, 76. DOI: 10.1007/s10462-024-10708-3.

[10] Steve Lydenberg, Jean Rogers, and David Wood. 2010. From Transparency to Performance: Industry-Based Sustainability Reporting on Key Issues. Tech. rep. Available from https://iri.hks.harvard.edu/links/transparency-performance-industry-based-sustainability-reporting-key-issues. Hauser Center for Nonprofit Organizations at Harvard University.

[11] M. Marzook and B. Al Ahmady. 2022. Linking organisational performance and corporate social responsibility. *European Jnl. of Business and Management Research*, 7, 335–343, 3. https://doi.org/10.24018/ejbmr.2022.7.3.1466.

[12] T. Nugent, N. Stelea, and J. L. Leidner. 2020. Detecting ESG topics using domain-specific language models and data augmentation approaches. (2020). http://arxiv.org/abs/2010.08319.

[13] Matthew Purver, Matej Martinc, Riste Ichev, Igor Lončarski, Katarina Sitar Šuštar, Aljoša Valentinčič, and Senja Pollak. 2022. Tracking changes in ESG representation: initial investigations in UK annual reports. In *Proceedings of the First Computing Social Responsibility Workshop within the 13th Language Resources and Evaluation Conference.* Mingyu Wan and Chu-Ren Huang, editors. Marseille, France, (June 2022), 9–14. https://aclanthology.org/2022.csrnlp-1.2.

[14] W. Xu and K. Eguchi. 2021. Topic embedding regression model and its application to financial texts. In *Proceedings of the Third Workshop on Financial Technology and Natural Language Processing*, 15–21.