Research in International Business and Finance 66 (2023) 102045

Contents lists available at ScienceDirect

Research in International Business and Finance

journal homepage: www.elsevier.com/locate/ribaf

ESG in the financial industry: What matters for rating analysts?

Marco Mandas *, Oumaima Lahmar, Luca Piras, Riccardo De Lisa

Department of Economics and Business, University of Cagliari, Viale S. Ignazio 17, 09124 Cagliari, Italy

ARTICLE INFO

JEL Classification:
G20
G24
Q58

Keywords:
ESG
ESG rating analyst
Banking
Topic modelling
Textual analysis
Sentiment analysis

ABSTRACT

This paper examines ESG rating analysts' views from Sustainalytics in order to highlight the main ESG features discussed across 11 sectors. We perform a topic modeling and a sentiment analysis to identify the content of analysts' opinions on the companies' ESG performance and to uncover the embedded sentiment associated with each ESG feature. The results of the topic modeling consist of 13 topics with a sector driven distribution. The analysis suggests that the best ESG performing financial institutions show to be actively committed to the code of best practice in governance and disclosure transparency. Whereas penalized financial entities seem to manifest less attention to ethical conduct and mis-selling. Furthermore, data privacy and security attract analysts' attention and should be closely monitored by financial entities. Finally, it is important to actively disclose ESG activities as the more information is available the better ESG commitment is reflected in analysts' views.

1. Introduction

Environmental, Social, and Governance (ESG) compliance is an emerging stream of research that recently gained traction due to the importance of these issues in the worldwide debate. In today's rapidly changing global landscape, the integration of ESG factors has emerged as a crucial aspect within the financial environment. Hence, a large and diverse set of standards, frameworks, and metrics related to ESG has been developed by regulators and policymakers around the world. Recently, several countries have introduced a range of different regulations. While some of these regulations provide businesses with more flexibility to develop their own best practices, others impose more stringent compulsory requirements (Jackson et al., 2020). Indeed, corporations and financial institutions are incentivized to keep track of their environmental and social footprint, and to disclose and report their activities. As a result, the percentage of firms that are publishing non-financial disclosures has increased massively from 35% in 2010 to 86% in 2021 (Rouen et al., 2022).

ESG considerations encompass a broad range of sustainability-related issues that go beyond traditional financial metrics. Institutions' ESG performance is becoming a growing focal point in financial markets, where investors decide and manage their investments also based on ESG criteria (Riedl and Smeets, 2017; Amel-Zadeh and Serafeim, 2018; Aureli et al., 2020). The availability of ESG ratings issued by rating agencies or, more generally, by specialized providers has become a crucial factor in mitigating information asymmetries and facilitating investors' portfolio choices (Brooks and Oikonomou, 2018). Furthermore, the growing recognition of the materiality of ESG issues by investors and financial institutions has resulted in the development of ESG performance assessment (Madison and Schiehll, 2021). These tools are intended to systematically evaluate, and benchmark companies based on their ESG performance, allowing investors to make more informed investment decisions aligned with their financial goals and sustainability

* Corresponding author.
E-mail address: marco.mandas@unica.it (M. Mandas).

https://doi.org/10.1016/j.ribaf.2023.102045
Received 26 January 2023; Received in revised form 14 July 2023; Accepted 22 July 2023
Available online 26 July 2023

objectives.

Financial and non-financial textual information is becoming with great importance for stakeholders as they provide, along with the quantitative data, a more comprehensive portrait of a subject company (Li, 2008; Rjiba et al., 2021). In fact, stakeholders are seldom interested in information especially if developed by professional analysts as they may be more knowledgeable, with higher expertise and less prone to biases (Bradshaw et al., 2017; Bradshaw, 2009). Furthermore, analysts' reports and recommendations are useful for the management as they may highlight ESG positive performances that should be reinforced or problematic aspects that need to be improved.

In this paper, we approach the matter by investigating the way rating analysts form, express, and discuss ESG ratings and the main issues highlighted by the analysts for the best and the worst financial entities in terms of ESG performance. The contribution of this work is, in fact, to shed the light on the issues that matter for ESG rating analysts, comments and suggestions provided by their reports, that may help guide subject companies in their management, improve their ESG engagement and better disclose their activities.

We analyze a database issued by Sustainalytics, containing ESG ratings of more than 5000 companies from 11 economic sectors. The Database contains not only the ESG scores of each entity but also the rating statements, in which the analysts explain textually the main reasons for the assigned score and other useful elements to understand the ESG performance.

Our approach is based on textual analysis (Loughran and McDonald, 2020), including two different methodologies recently used in academic research: topic modeling and sentiment analysis. They are based on computational operations applied to textual data that combine linguistics, semantics, computer science, and statistics to uncover latent information and sentiments in a given text (Lin and He, 2009).

We apply topic modeling, a machine learning technique originally developed by Blei et al. (2003), to companies' rating statements to investigate the content of analysts' views on companies' ESG performance. This methodology allows to identify and cleanse the dominant topics concerning ESG integration assessment, which, then makes it possible to sort companies based on their ESG performance expressed in terms of scores and analysts' judgment. We also perform a sentiment analysis or opinion mining, a Natural Language Processing (NLP) technique, by which the analysts' opinion is analyzed to uncover the sentiment embedded in the text, whether positive or negative (Hu and Liu, 2004). The objective of the latter is to assess if the texts are more likely to stimulate the readers' positive or negative feeling. Merging the outputs of both methodologies provides a more comprehensive framework which paves the way to a different approach in studying and interpreting texts. Textual data contains a wide set of latent information seldom undetectable or maybe neglected, though having great revealing potential, impossible to investigate otherwise.

Our results suggest that analysts' ESG views are well featured by 13 main topics that combine differently across economic sectors and highlight significant sectoral diversity of ESG issues. Furthermore, focusing on analysts' views about financial institutions reveal sector specific features where the attention is driven to a greater extent on governance compared to the other issues (Ethical conduct, ESG integration in the financing decisions, Product governance transparency).

After sorting out the best and the worst performing companies in terms of ESG scores, we show that the best-scored ones take greater care of data privacy and product governance transparency. On the other hand, analysts call attention to the ethical and misselling conduct held by the worst financial institutions in terms of ESG performance. Thus, the results suggest that the more transparent the financial institution is, the more prone it is to reach a higher ESG rating (Forssbaeck and Oxelheim, 2014). This study brings insightful contributions to the available literature on the matter. In fact, using textual analysis integrated with quantitative information features the complementarity between qualitative and numerical elements. In addition, we provide the industry with practical evidence about ESG issues subject to analysts' attention and assessment, that may be with great support to their managerial activities in dealing with ESG concerns.

The first section presents an extended literature review focusing on ESG-related studies performed using textual analysis and specifically NLP tools. The second section describes the data and the methodology adopted to perform the analysis. Then, the results are discussed in the third section including the major implication of the research's findings. The last section is the conclusive part where we expose the main contributions of the paper.

2. Literature review

The increasing relevance and importance of ESG practices for governments, investors, and stakeholders called the academic community for extensive, renewed, and deeper investigations. This can be observed through the growing number of publications about ESG reports and the methodological diversity applied to examine them. Over the past decades, the number of corporations disclosing detailed reports on their "sustainable" activities and initiatives has been rapidly increasing (Benvenuto et al., 2023). It comes with no surprise that the focus is growing on the added value of the textual information provided by the disclosing entity, which generates a research opportunity.

Textual information may embed latent contents such as concerns, sentiments, doubts, qualified opinions, weaknesses, threat perceptions, consensus drivers, and many other important features. The latter are used for the design and implementation of managerial actions, policy strategies, and assessments. All these would lead to proper conjunction between "reality" and the "sense of reality", typical of the language, as Bertrand Russel would have put it (Russell, 1959). In other terms, the large existing literature on textual analysis and its effects on human decision processes, though supported by rigorous methodologies, still leaves some room to further speculation in order to shed the light on the mechanisms through which it affects human actions (Barberis et al., 1998).

Textual communications proved to be powerful and influential on decisions, sentiments, perceptions, and risk preferences, which may explain the increase in the number of publications applying empirical methods on finance-related documents. Studies performing textual analysis focused on annual disclosures as a source of information addressed to stakeholders, give new interpretation keys and a more comprehensive overview on corporate financial performance (Li, 2008). Meanwhile, with the growing interest in corporate sustainability performance, ESG disclosures as well became subject of investigation in the academic literature (Hahn and Kühnen, 2013). As a matter of fact, these methods are applied to dig deeper into the qualitative data to obtain additional information, not only about the financial status of corporations, but also to evaluate their commitment to ESG criteria (Baier et al., 2020).

Content analysis on corporate communications has been conducted to examine the nature and the degree of disclosure (Deegan et al., 2002), the influence of media on the information disclosed (Baier et al., 2020), and even to track changes in corporate releases over time (Jenkins and Yakovleva, 2006). By establishing ESG checklists, word lists or dictionaries, and categories, it is possible to score the disclosing entity and evaluate how mature it is in terms of ESG compliance (Deegan et al., 2002; Jenkins and Yakovleva, 2006; Giles and Murphy, 2016; Baier et al., 2020). This analysis is based on setting or listing ESG criteria and determining their related words or dictionaries. Through checklists and word frequencies in a document, it is possible to understand the importance of a particular criterion and the attention it gets from the disclosing entity or stakeholders.

However, manually navigating ESG reports seems to be a daunting, time-consuming, and bias-prone task which also results in a smaller number of examined data. More recently developed techniques such as machine learning models applied to textual data facilitated the extraction of information from large corpora, which became a less complicated task, more reliable, and less time-consuming.

Topic modeling and sentiment analysis are recently developed methodologies by which it is possible to obtain relevant information present in a textual document. Both methodologies use NLP and computational tools on textual data. Topic modeling permits to extract latent themes (Blei et al., 2003) whereas sentiment analysis aims at extracting opinions and sentiments (Liu, 2020). The literature on topic modeling and sentiment analysis in finance and more specifically on ESG-related texts is still nascent. As far as we are concerned, only few papers used these approaches on ESG data.

In a number of papers, topic modeling has been applied on ESG reports where authors set different research objectives. Ning et al. (2021) tried to identify the users of ESG reports based on the topics disclosed and their targeted readers. The authors claim that the focus on ESG topics and themes to which a corporation is committed may bring financial value. The paper also provides an insightful finding about the strategic corporate intent of managing external relationships with stakeholders through voluntary disclosure on ESG. In the same context, Lindgren et al. (2021) identify the targeted readers of ESG reports based on the nature of the topics disclosed. The findings confirm the claim of Ning et al. (2021), as firms adopt a stakeholder approach to management where they are favored.

Another line of papers uses voluntary corporate disclosure to identify the main ESG themes and topics described by the latter and examine their distribution among different sectors. Székely and vom Brocke (2017) extracted forty-two topics from 9500 corporate sustainability reports, that are equally distributed among environmental, social, economic, and general sustainability themes. Topics exposed under environmental issues are argued to be particularly linked to emissions and energy consumption. Parallelly, topics disclosed as a part of social sustainability mainly refer to employees' well-being and customer orientation. Furthermore, Golosh-chapova et al. (2019) find similar results and show a sector bias in the distribution of the topics among the major three ESG areas, both in the EU countries and the UK.

ESG-related information is not exclusive to sustainability reports, in fact, this type of information is also present in traditional and social media sources, where textual data can be gathered from disclosing entities as well as users and stakeholders. In fact, Lee et al. (2023) used topic modelling across discourse sources on ESG. Particularly, based on the topics' evolution over time in the academic debate and in media, the authors found that the value of utility was first debated in academia following the schemes of the United Nations organization, and then the media focused on these issues. Meanwhile, Park et al. (2022) use 73,397,870 text data scraped and cleaned from Twitter, to track the evolution of ESG related topics using dynamic topic modeling. The authors extracted 18 topics that capture the ESG debate on twitter, which they further classify into four major themes: environment, investment, education, and evaluation. Regarding ESG debate, topics discussed and sentiments expressed show an evolution overtime that translates the public opinion which in turn has been examined by Liu et al. (2023). In both papers, the authors combine descriptive analysis, topic modeling, and sentiment analysis to examine users' characteristics and interests regarding ESG issues. Liu et al. (2023) reveal that, despite the increasing number of Weibo posts on ESG, there hasn't been a proportional increase in attention received by the platform users. The comments expressed positive, neutral, and negative sentiments at rates of 66.03%, 17.37%, and 16.60% respectively. Negative sentiments are likely influenced by factors like greenwashing practices, limited knowledge about ESG, inconsistencies in ESG ratings, and lack of transparency in rating methods.

Likewise, Shen and Li (2023) identify several topics discussed in social media networks related to green housing, showing that negative sentiments are mainly caused by price related issues. In the same context, sentiment analysis shows to have an important role in determining the sentiments associated with ESG issues expressed by both corporations and the general public. Bapat et al. (2022) investigates the market reaction to ESG news revealing an association between the sentiment of market news and stock prices. Similarly, Serafeim and Yoon (2022) examine if and how investors respond to the sentiment transmitted through ESG news. The authors show that not all ESG-related events covered by the news have a significant price reaction.

To further understand the factors influencing corporate commitment to ESG disclosure, Amin et al. (2021) identify ESG matters publicly shared on social media and show that there are mainly three factors positively associated with the extent to which corporations are engaged to ESG public releases on twitter: board gender diversity, independence and company size. In fact, corporate engagement in constant ESG press releases contributed to a lively debate on the matter and resulted in new practices in the field such as the introduction of ESG environmental management, spread of ESG environment-friendly management, and implementation of ESG carbon neutrality (Seo et al., 2022).

Analysts prove also to play a significant role in the matter, and, in fact, their work including recommendations, evaluations and ratings is frequently considered in the debate. One important stream of literature goes back to two decades and covers a wide range of biases that affect analysts' reports depending on their level of independence from either the information provider or the target company (Lang and Lundholm, 1996). More recently their narrative has been an interesting object of inquiry. Kiesel and Lücke (2019) examine ESG consideration in rating reports published by credit rating agencies. The authors study the stock returns and credit default swap (CDS) spread changes to check whether ESG consideration has an effect on the capital market reactions. They show a minor, but existing consideration of ESG in rating decisions. Corporate governance holds the utmost significance within the realm of ESG considerations. Additionally, the findings indicate that the incorporation of ESG factors significantly impacts stock returns and credit default swap (CDS) spreads when rating announcements are made.

In the light of the discussed literature, the scarcity of studies examining ESG rating agencies, calls for further investigations of the matter. There is a wide range of articles studying the ESG evaluation tools, measurements, and scores, however, textual comments that come along numerical scores provided by rating agencies seem to be overlooked if not totally absent from the debate. Given the importance of analysts' opinions, evaluations, and comments, not only for managerial matters, but also for stakeholders and investors, it is crucial to perform a comprehensive analysis that combines both numerical and textual ESG evaluations. The present investigation makes a three-fold contribution to the debate on ESG. First, we combine novel tools, topic modeling and sentiment analysis, to extract latent information embedded in ESG ratings analysts' textual comments. Hence, we provide managers with deeper practical insights about what matters for ESG evaluations. Second, we integrate qualitative and quantitative inputs in a unique framework, that offers stakeholders a comprehensive overview on the most influential factors affecting ESG performance across sectors, and specifically in the financial industry. Third, we bring to the debate a new type of textual information, ESG rating analysts' opinions, that, to our knowledge, has not been studied so far.

3. Data and methodology

3.1. Data

The paper uses the ESG Risk Rating Focus Database developed by Sustainalytics, a Morningstar company.[1] The database contains more than 100 variables of ESG performance for 12,480 entities, covering a time span starting from November 2020 to January 2021. It provides several ESG scores, measuring selected issues presenting material risks to entities' performance. Overall, scores are provided with qualitative (textual) assessments, explaining companies' ESG performance and their capacity to manage related risks. These assessments represent the judgement for three material factors considered by the analyst to be of high priority for each company. Each analyst's view forms a commentary on the company's performance with respect to the exposure and the management of the selected material ESG issue. On average, analyst views' length is about 200 words each..

In this paper, we select the entities having the three (qualitative) analyst's views and the overall (quantitative) ESG risk score. This leads to an ultimate selection of 5012 entities, out of 12,480, for Dec-2020[2] (Table 1). Companies belong to 11 sectors; the financial sector counts for 19% of the total entities. Considering continents breakdown, Europe has a weight of 31%, Asia/Pacific 35% and US 25%.

3.2. Methodology

The ESG Risk Rating Focus Database provides unique data related to ESG risk through both numerical scores and textual comments developed by ESG analysts, that to our knowledge, hasn't been examined in the available literature. To better explore the data, it is necessary to combine both textual and empirical analysis instruments. Topic modeling seems to be significantly informative as it is a text mining tool aiming at detecting and discovering latent topics discussed in a corpus, which is in this paper, the analysts' views. Each topic is represented by the most frequent word clusters that co-occur in a set of given documents based on which it provides a set of topics or themes dominating the input corpus. As a result, the output defines the major issues discussed by rating analysts(topics). In this context, we applied the Latent Dirichlet Allocation (LDA) approach, which is one of the many generative statistical models used in topic modeling. It is defined as a dimensionality reduction technique attempting to discover the latent thematic structures in a corpus (Blei et al., 2003).

Furthermore, we apply sentiment analysis (Hu and Liu, 2004) along with topic modeling to understand the analysts' sentiment associated with certain topics and better understand what ESG issues could potentially provoke positive, negative or no sentiment at all. The combination of both methodologies is an interesting research approach to textual documents, especially in the context of ESG given that they, combined, provide an in-depth investigation tool (Amin et al., 2021; Liu et al., 2023). In addition, we use the numerical scores to further understand how companies are categorized by associating ESG scores to the analysts' sentiments.

The analysis follows a three-step process. The first step consists of applying topic modeling to the entire set of analysts' views related to entities in all sectors in the dataset. The aim is to extract an acceptable number of topics from the qualitative assessments made by the analysts. Secondly, we focus on the financial sector by exploring in depth the distributions of each extracted topic and the

[1] Sustainalytics reports that "ESG Risk Rating captures an issuer's exposure to the material, industry specific ESG risks and an issuer's management of those risks. The ESG Risk Rating provides investors with an overall company score based on an assessment of how much of a company's exposure to ESG risk is unmanaged. The more of this risk that is unmanaged, the higher the ESG Risk Rating score".

[2] We select the textual data referred to a single period of time (Dec-2020) as the analysts' qualitative assessment do not change over the short time span that is covered in the database (from Nov- 2020 to Jan-2021)

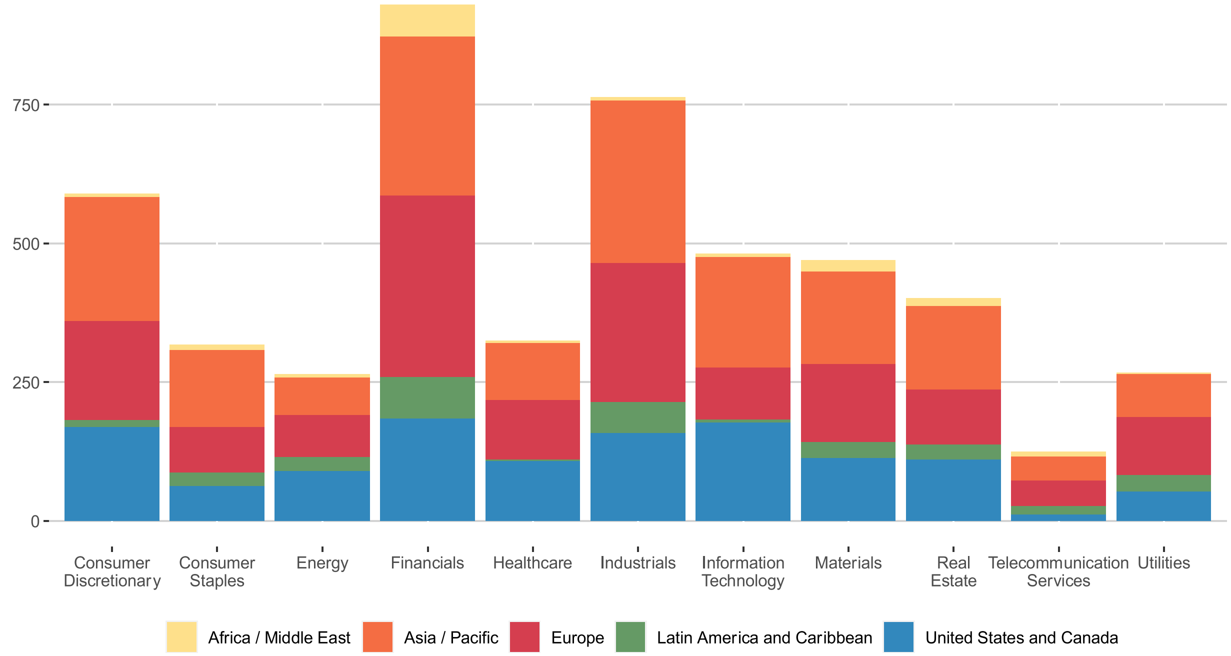

Chart 1. Breakdown of the entities in the database divided by sector and origin (colours). The y-axes reports the absolute number of the entities from the corresponding sector.

different features discussed by the analysts. Finally, a sentiment analysis is performed on analysts' views to uncover the valence of the embedded sentiments related to the best and worst ESG performing financial entities.

In order to obtain reliable analytical results from topic modeling, it is necessary to preprocess the textual data. The preprocessing stage ensures that only relevant and informative textual features are retained for the machine to analyze. The paper adopts the following processing sequence by using RStudio to prepare the corpus: i) cleaning the text which involves different tasks (switch to lowercase, removing special characters, symbols, numbers, and punctuation); ii) removing stop words such as "is", "the", "and" by using the list provided by topic models package in R; iii) tokenizing the text; iv) lemmatizing the words (Asmussen and Møller, 2019).

Finally, we also remove some words that are extremely common in the analysts' views: a list of 42 words[3] that occur in more than 25% of all the documents. These words have limited discriminatory power and do not help in classifying topics, preventing the model from estimating them properly. Additionally, we also exclude all words that do not occur in at least 10 documents. Quite the opposite, these words have too high discriminatory power as they may refer to a specific company or a particular item and, consequently, are not useful for topic analysis.

Another methodological challenge is about the choice of the number of topics (K) to be extracted. Unlike other quantitative approaches, LDA is an automated machine learning tool, nevertheless, it requires human intervention to work properly, namely the number of topics has to be set in advance. Estimating the number of topics in advance is a challenge in LDA since users do not have prior knowledge of the exact number of topics present in the corpus.

For this reason, we use the perplexity score to guide the choice of K. It is virtually the most used score in the literature given the simplicity of its interpretation (Zhao et al., 2015). In fact, the perplexity is a statistical measure that is used to compare the generalization performance of models with different values of K. The score is inversely proportional to the geometric mean per-word likelihood. The higher is the geometric mean likelihood, the lower is the perplexity score and, eventually, the better is the generalization performance of the model (Blei et al., 2003). In order to interpret the evolution of the perplexity score and identify the optimal number of clusters, we use the "Elbow rule". The goal is to find the number of topics where adding more topics does not significantly reduce the perplexity score. In other terms, we identify the "elbow" of the curve, where the decrease in the perplexity score slows down noticeably, resulting in a significant change in the slope of the curve. As shown in Chart 2, for $K = 12$ the perplexity score decreases abruptly and then it starts to decrease at a slower rate, which implies that K is between 12 and 15.

Given that the perplexity score does not provide a precise number of clusters, we opt for a trial error method by which we set K to 12, 13, 14 and 15, and examine the resulting word clusters in terms of precision, coherence, and clarity. After analyzing the results associated with a range of values of K, we decide to set the parameter to 13 topics, considering the quality of the output and the degree

[3] The list of words that have been removed below: company, issue, average, management, exposure, risk, view, subindustry, medium, include, similar, practice, relate, low, product, program, address, moderately, measure, controversy, service, implicate, preparedness, industry, business, system, increase, market, employee, strong, customer, impact, cost, minor, expose, operation, lead, result, policy, regulatory, report, train.

Table 1
Breakdown of entities by sector and origin. The table reports the absolute value of the entities from a specific sector and region, the percentage with respect to the total number of entities (column called “total”) and with respect to the total number of entities from the corresponding area (column called “area”).

Sector	Africa / Middle East	%		Asia / Pacific	%		Europe	%		Latin America and Caribbean	%		United States and Canada	%		Total	%
		total	*Area*		*total*	*Area*		*total*	*Area*		*total*	*Area*		*Total*	*Area*		
Consumer Discretionary	7	1.18%	4.70%	223	37.7%	12.7%	181	30.6%	11.7%	12	2.0%	3.8%	169	28.5%	13.6%	592	12%
Consumer Staples	10	3.11%	6.71%	139	43.2%	7.9%	86	26.7%	5.5%	24	7.5%	7.7%	63	19.6%	5.1%	322	6%
Energy	7	2.62%	4.70%	68	25.5%	3.9%	76	28.5%	4.9%	26	9.7%	8.3%	90	33.7%	7.3%	267	5%
Financials	59	6.12%	39.60%	290	30.1%	16.5%	348	36.1%	22.4%	82	8.5%	26.2%	185	19.2%	14.9%	964	19%
Healthcare	5	1.53%	3.36%	102	31.2%	5.8%	110	33.6%	7.1%	2	0.6%	0.6%	108	33.0%	8.7%	327	7%
Industrials	7	0.91%	4.70%	294	38.1%	16.7%	255	33.1%	16.4%	57	7.4%	18.2%	158	20.5%	12.7%	771	15%
Information Technology	6	1.24%	4.03%	199	41.2%	11.3%	95	19.7%	6.1%	5	1.0%	1.6%	178	36.9%	14.4%	483	10%
Materials	21	4.43%	14.09%	166	35.0%	9.4%	143	30.2%	9.2%	29	6.1%	9.3%	115	24.3%	9.3%	474	9%
Real Estate	14	3.43%	9.40%	153	37.5%	8.7%	102	25.0%	6.6%	29	7.1%	9.3%	110	27.0%	8.9%	408	8%
Telecommunication Services	9	7.09%	6.04%	44	34.6%	2.5%	47	37.0%	3.0%	16	12.6%	5.1%	11	8.7%	0.9%	127	3%
Utilities	4	1.44%	2.68%	79	28.5%	4.5%	110	39.7%	7.1%	31	11.2%	9.9%	53	19.1%	4.3%	277	6%
Total	149		2.97%	1757		35.1%	1553		31.0%	313		6.2%	1240		24.7%	**5012**	100%

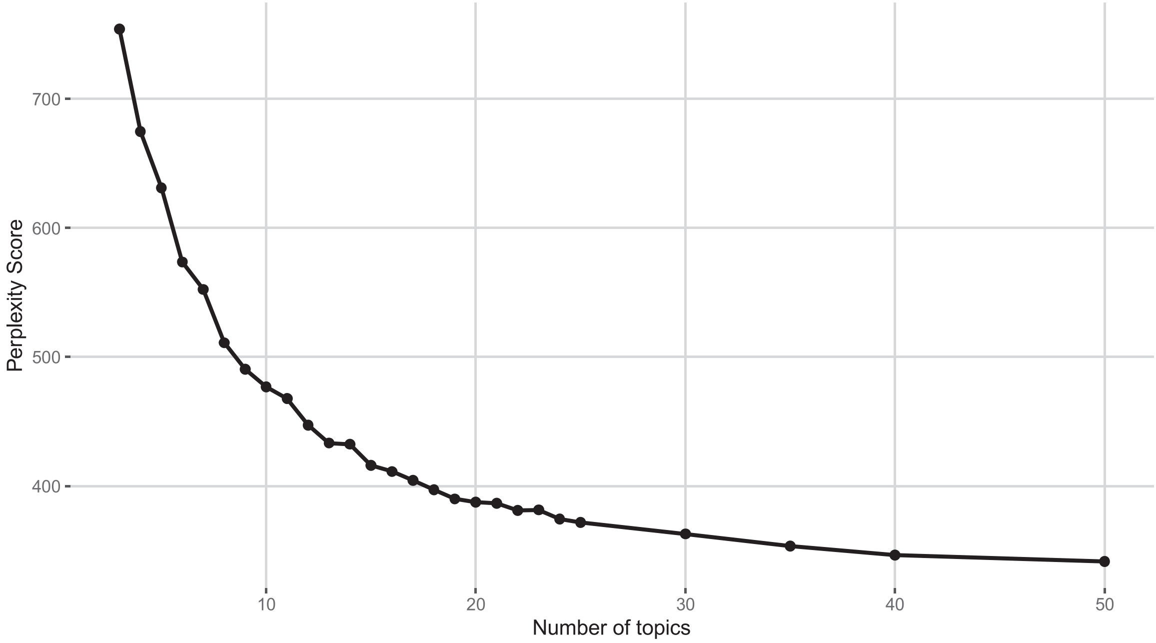

Chart 2. Perplexity score.

of topic precision the model provided with K = 13.

Once estimated the LDA model, the results are 13 clusters of words that represent the dominant topics or, in our context, ESG issues discussed by rating analysts. To facilitate the analysis further performed and better analyze the ESG features that matter, it is necessary to set descriptive labels for each topic. Labelling word clusters is far from being a trivial step or a simple formality. Labels denote the topics' content and synthetically define the semantic field within which a given aspect of analysts' view is addressed. The process starts by inspecting the most relevant words in every word cluster according to the λ (*lambda*) measure[4] preset to 0.5, provided by *LDAvis* application available in R. Then, the analysts' views that present the highest probability to belong to a particular topic[5] are identified in order to perform a concordance analysis by which each relevant word is inspected in its original context. The extracted inputs are accurately examined and contextualized to choose the most suitable label. By exploring the most relevant terms' concordances within the analysts' opinion, the final labels are set.

The second phase of the investigation aims at uncovering the factors analysts focus on the most, when assessing the ESG performance of the Best and the Worst financial entities. It consists of an in-depth analysis of two subsamples of entities, strictly belonging to the financial sector, comprising the "Best" and "Worst" financial institutions in terms of ESG scores respectively, in order to dig deeper into strengths and weaknesses that matter for the ESG analysts. To this end, the entities with the highest ESG scores (upper 25% percentile) and those with the lowest scores (lower 25% percentile) are identified. On these two sub-samples, the most frequent topics are closely analyzed.

The idea behind the previously described steps is to provide a more precise framework to perform the sentiment analysis. In fact, the third and final task is aimed at identifying the analysts' sentiment implied in their judgments on the best and worst financial entities. Sentiment analysis is useful to understand whether the previously highlighted ESG issues are discussed by analysts with a positive or negative sentiment. It is, at this point, possible to find the most important ESG factors, associated with positive or negative sentiments, that the analysts express when commenting the ESG commitment of the best and worst performing financial entities.

In order to determine the conspicuous sentiment, we use the polarity function in R, that has a built-in subjectivity lexicon[6] in which each word is assigned a particular sentiment value. In order to score analysts' opinions, the function scans the text and matches the words in the input document with those in the lexicon and assign the corresponding score. The function also takes into account the intensity of the sentiment by creating a cluster around an identified word, where valence shifters adjust the score. Valence shifters are words that amplify or negate the emotional intent of the word. For example, "well known" is positive while "not well known" is negative. Here "not" is a negating term and reverses the emotional intent of "well known". By contrast, "very well known" employs an

[4] λ is the weight given to the probability of a term within a topic relative to the ratio of that term's probability within a topic to its marginal probability across the corpus. The higher is λ, the higher is the weight given to the most frequent terms within a topic. The lower is λ, the higher is the weight given to the terms presenting a greater discrepancy between the frequency within a topic and the frequency within the corpus. We set λ = 0.5 that is considered an optimal value to compute the "relevance" of the terms.

[5] This information is included in the Gamma matrix in the LDA results.

[6] It is a predefined list of words associated with emotions or positive/negative feelings.

amplifier, increasing the positive intent. The polarity function aggregates the scores by calculating the overall sentiment based on subjectivity terms, valence shifters, and the total number of words.

There are terms like "issue" and "risk" that are widely used in the analysts' views that tend to lower the sentiment score. As a result, 80% of the views have a negative sentiment score. Considering that such negative terms seem to be used in the same proportion for each view, we could assume that the sentiment scores incorporate a constant negative bias. For this reason, we normalize the scores to be between 0 and 1 by using the following formula: $z_i = (x_i - min(x)) / (max(x) - min(x))$. Then we evaluate if the normalized values are above or below the average.

4. Results

The first results illustrated in Table 2 consist of 13 extracted topics from the corpus. The output of the topic modeling is a set of keywords that have the maximum likelihood to represent each topic. Table 2 reports the top 10 words that probabilistically best describe the topics. Thus, based on the most frequent word pattern of each topic and the concordances they have inside the corpus, the topics are labeled. The labeling, as mentioned above and according to the literature is a human guided activity. In other terms, it is essential that the researchers rigorously combine the analysis outcomes with their knowledge, logic and even intuition to enhance the topic label accuracy. After labeling the topics, we categorize them into three major groups: three topics related to environmental issues, five topics associated with social impacts and five governance-related topics that analysts express in their views.

To further understand whether the financial sector shows different ESG features compared to the other sectors, namely how the major three categories are distributed among them, we track the proportions of each topic in each sector as shown in Chart 3.

The emerging finding from Chart 3 is the increasing proportion of governance-related topics when approaching sectors characterized by substantial governance requirements, such as the Real Estate and Financials. On the other hand, environmental issues are given less attention in the latter sectors contrarily to industries like Energy, Materials, Consumer Staples, and Utilities, where they are strongly featured. Meanwhile, the social-related topics tend to exhibit a stable presence in analysts' views across all activity sectors with a dominant appearance in Healthcare and IT sectors.

Overall, the distribution of the topics is clearly sector driven (Gillan et al., 2021; Goloshchapova et al., 2019; Jayarathna et al., 2021). In fact, rating analysts evaluate the performance of each subject company within its activity sector and guide their judgment based on the ESG area the company is influencing most whether positively or negatively. Sectors face obvious differences in aspects that are relevant or critical. Such differences are mostly related to the dominant business model within the sector, and they tend to manifest them in aspects like competitive pressure, supply-chain characteristics, customer/market burdens and general environment ones. To eventually stress further this idea, one can jointly consider sector's key activities and the major business model factors, mainly the way these businesses generate revenues and support costs (Teece, 2010), in order to provide a more precise framework for the ESG features emerging from the analysis (See Appendix B).

It is not surprising that the three environment-related topics analysts discuss in input corpus are mainly associated with Energy, Utilities and Materials. These are sectors characterized by a potential significant environmental footprint, given the key activities and

Table 2
Topics labels.

#	ESG	Label and most frequent words
1.	E	**Sustainable energy** *Carbon, Energy, Emission, GHG, Reduce, Regulation, Renewable, Fuel, Power, Gas*
2.	E	**Environmental and climate change risk** *Emission, Waste, Community, Effluent, Environmental, Air, Project, Relation, Mine, Spill*
3.	E	**Sustainable supply chain** *Water, Supply, Resource, Right, Chain, Supplier, Material, Human, Production, Manufacture*
4.	S	**Human capital management** *Human Capital, Diversity, Manage, Initiative, Support, Workforce, Disclose, Growth, Recruitment, Skill*
5.	S	**Human capital empowerment** *Human Capital, Labour, Development, Initiative, Turnover, Train, Talent, Diversity, Rate, Develop, Experience*
6.	S	**Data privacy and security** *Data, Security, Privacy, Information, Breach, Process, Cybersecurity, Sensitive, Financial, Cyberattacks*
7.	S	**Consumer protection** *Quality, Governance, Safety, Standard, Recall, Test, Ensure, Certification, Food, ISO*
8.	S	**Health and safety** *Safety, Health, Occupational, Fatality, Incident, Contractor, Train, Accident, Performance, Injury*
9.	G	**Business ethics and code of conduct** *Ethic, Corruption, Bribery, Ethical, Whistleblower, Anticompetitive, Bank, Fine, Tax, Launder*
10.	G	**Product governance transparency** *Disclose, Conduct, Major, Commitment, Access, Potential, Train, Lack, Initiative, Strategy*
11.	G	**ESG integration in investment and financing decision** *Integration, Investment, Financials, Bank, Responsible, Opportunity, Asset, Project, Diversify, Finance*
12.	G	**Real estate green finance** *Build, Property, Real Estate, Revenue, Tenant, Green, Safety, Portfolio, Climate, Building*
13.	G	**Ethics conduct and misselling** *Governance, Quality, Financial, Client, Bank, Offer, Lend, Control, Reputational, Poor*

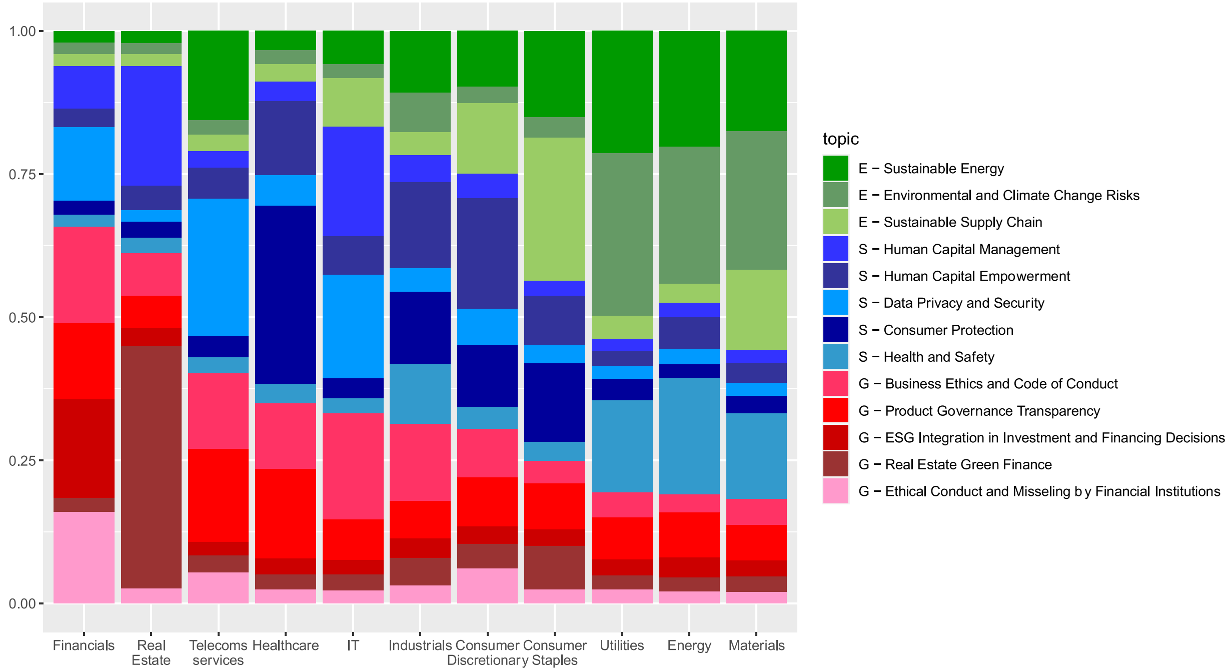

Chart 3. Distribution of topics by sector.

the main sources of revenues and costs, all of them representing solid ground for opportunistic behaviors conflicting with ESG guidelines. Particularly, the analysts highlight features related to sustainable supply chain for the two consumer-oriented sectors (Consumer Discretionary and Consumer Staples). On the other hand, the topics more related to carbon emissions, greenhouse gas production, and community integration ("Sustainable Energy" and "Environmental and Climate Change Risks") are prevalent in the sectors of Energy, Materials, and Utilities.

Focusing on the financial sector, topics related to environmental issues have a marginal weight, minimal when compared to other sectors, in analysts' views. Indeed, financial institutions play a negligible role in directly affecting environmental aspects. However, they are crucial in determining which criteria are relevant to give access to financing, which is significantly featured in the governance-related topics about ESG integration. In fact, He et al. (2021) argue that loans initiated after an ESG related incident have significantly higher spread and lower loan size. Likewise, Chen et al. (2023) find that ESG performance can directly and indirectly reduce the cost of equity capital. Hence, it can be suggested that financial institutions have not only a direct role in managing their own ESG performance, but they have also an indirect effect translated by whether to grant access to financing to other firms based on their ESG performance.

Furthermore, unlike environment-related topics, governance issues attract analysts' attention when it concerns financial institutions. Governance aspects play a central role in financial institutions' evaluation as they are subject to rigorous regulatory requirements involving their responsibility in managing capital, maintaining financial stability, and monitoring financial risk. In addition, the fiduciary responsibility acts as a major factor influencing the way financial institutions deal with their duty to act in the best interest of their clients and shareholders. This implies ensuring proper corporate governance practices within the institution itself, such as executive compensation, board independence, and ethical conduct (de Haan and Vlahu, 2016). Meanwhile, investors are increasingly interested in ESG issues since they are considered as a significant indicator of long-term value and risk management. Thus, they expect from financial institutions to show strong governance practices associated with board composition, ownership structure, and risk management in general.

In fact, Chart 3 shows that the highest proportion of the extracted topics is governance related, featured by "Business ethics", "Product governance transparency", and "Ethical conduct and mis-selling". Precisely, analysts discuss moral aspects such as anti-corruption measures, commitment to transparency in public disclosures and compliance with the code of best practice. However, the debate in the literature on the impact of disclosure transparency on financial institutions' market value is still not conclusive as the public disclosure is not an economic activity. Bhimavarapu et al. (2022) claim that transparent disclosure does not impact banks' valuation, but it is positively associated with banks' market value under the influence of the overall ESG score. This insight is in line with the above shown results as the presence of governance transparency and disclosure among the topics discussed by analysts is significant given the proportion of these topics in the overall corpus. In a nutshell, "Product governance transparency", "Ethical conduct", and "Business Ethics", are mechanisms and strategies that may contribute to reduce information asymmetry and agency costs by which financial institutions specifically, and firms in different sectors generally, can potentially impact their firm value.

It is worth highlighting that "Data privacy" is a matter of interest not only in the two technology sectors (Information Technology and Telecommunication Services), but also in the financial one. The financial sector, in its commitment to embrace the technological advancements, is one of the first sectors applying artificial intelligence. However, this application is a double-edged sword. It improves fraud detection, money laundering, assists decision making processes and most importantly enhances the quality of customers assistance with more personalized services. Nevertheless, to reach these improvements, AI may also increase the risk of data breaches, which are highly sensitive in the financial environment. Thus, protecting sensitive data is increasingly important and analysts frequently warn financial institutions to pay attention to it.

More generally, social issues appear evenly distributed across all sectors. Differences are much less significant compared to environmental and governance ones. However, it is noteworthy that in Healthcare, IT, Industrials and Consumer Discretionary sectors, social-related issues attract more expressively analysts' attention. Furthermore, Human capital management and human capital-oriented policies are relevant for virtually all sectors. Yet, they exhibit a significant appearance in Industrials, Consumer Discretionary, IT, and Real Estate sectors. The explanation of this finding may be straightforward as these sectors are highly dependent on the quality of the work environment (Industrials), the growing qualification requirements (Healthcare), data privacy (IT) and the attention toward the respect of emerging social standards, leading to a sound ESG compliance.

In order to respond to the research question concerning the ESG aspects featured by the best and the worst performing financial institutions, the latter are divided based on their ESG scores in two subgroups with the highest and lowest ESG scores, that is the upper and the lower 25% percentile of the sample scores. Then, both subgroups are associated with the main topic, or, in this context, the major concern commented by the rating analyst.

Furthermore, the analysis includes the rating analysts' sentiments embedded in their views concerning both the best and worst performing institutions in terms of ESG. Hence, it is possible to provide a comprehensive investigation involving topics that matter for rating analysts, ESG features that characterize the best and worst performing entities and the prevalent sentiment in the analysts' views. Chart 4 shows the size and the sentiment valence of each topic for both clusters of financial institutions.

The first insight is that the analyst's views about the financial institutions with better ESG performance emphasize two key factors that contribute to positive assessment, the quality of data privacy and security systems and their commitment to transparency. As illustrated in Chart 4, the "Product Governance Transparency" and "Data Privacy and Security" topics are more relevant and are discussed by the analysts with a positive sentiment for the Best-scored institutions. This finding is in line with the insights of the

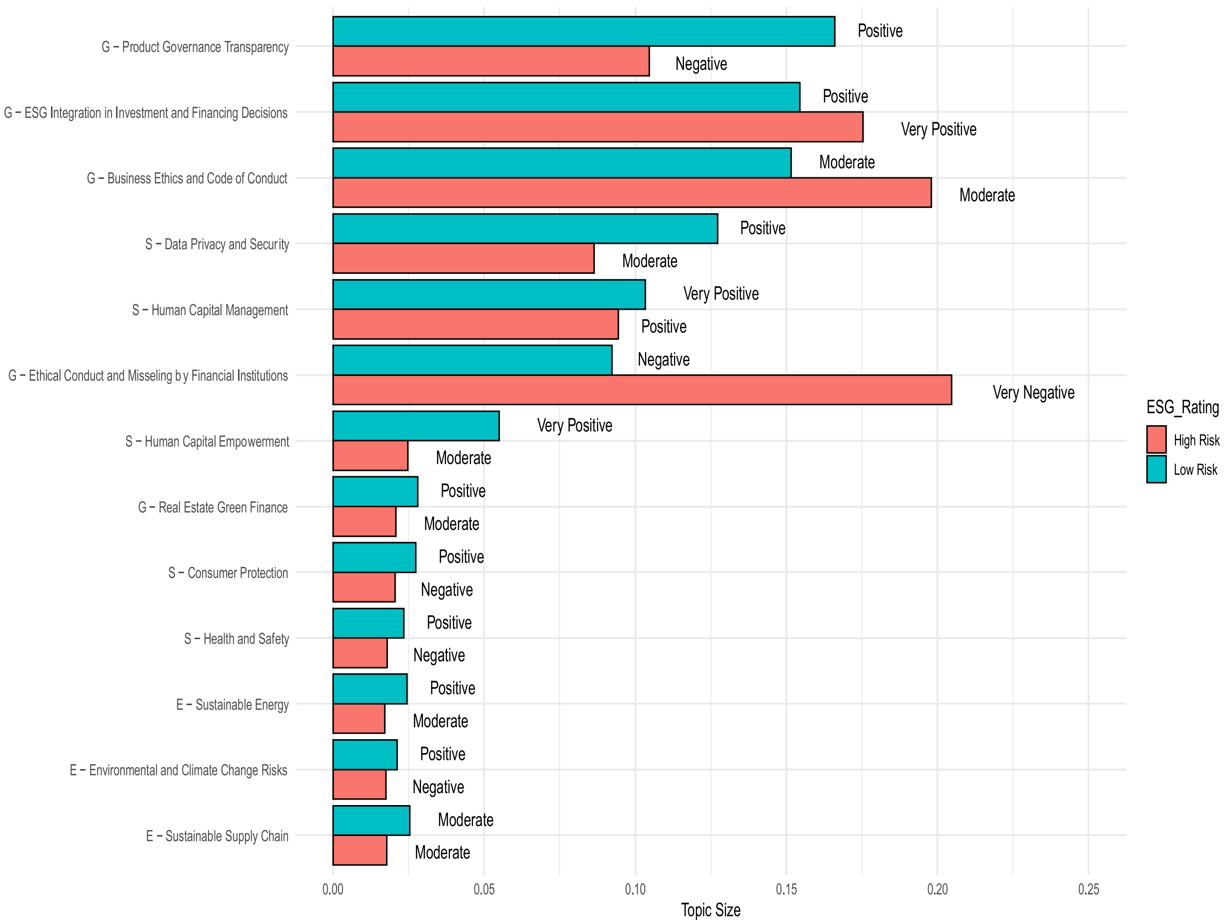

Chart 4. Proportion and distribution of topics between "Best" and "Worst" Financial Institutions where the labels next to each bar indicate the average sentiment score in a scale from "Very Negative" to "Very Positive".

previous discussion, highlighting the fact that data privacy and security is a growing concern for individuals. Analysts, through their positive sentiment, recognize that financial institutions with best ESG score implement robust privacy measures and safeguard clients' information from breaches. Hence, these entities demonstrate a socially responsible approach to managing sensitive data, instilling trust in their clients and stakeholders. Besides, governance transparency is as well associated with both high ESG scores and positive analysts' sentiment when financial institutions are seen as actively disclosing their managerial activities in a transparent way. By practicing transparent disclosure, financial entities foster confidence in their clients and stakeholders and demonstrate their commitment to accountability, which contributes to a better assessment of their ESG performance.

On the other hand, the second significant finding is that financial institutions exhibiting ethical issues and mis-selling conduct may be seriously penalized by ESG analysts. Chart 4 shows how analysts talk to a great extent about these topics ("Ethical conduct and mis-selling by financial institutions" and "Business ethics and code of conduct") with a very negative tone in reference to the worst financial entities. In fact, issues like discriminatory practices, involvement with controversial activities, exploitative behaviors and more generally morally unaccepted activities are viewed negatively by rating analysts, graphically reflected in the intensity of the associated negative sentiment that is significantly higher compared to negative sentiments associated with other ESG aspects. Consequently, it can be suggested that analysts acknowledge ethical conduct as a significantly important factor that when financial institutions fail to respect it or to act accordingly, are deemed to have a negative ESG assessment. Therefore, it would be highly recommended to limit actions with bad reputational consequences in order to get a higher ESG score.

These two results highlight how transparency is the way to increase the probability to obtain a good ESG score. Hidden strategies, unethical behaviors, and lack of transparency are not favorably welcomed by the analysts that tend to severely penalize them. This evidence may be attributed to the evaluation procedures employed to assess ESG performance, which heavily rely on public disclosures, news, and non-governmental reports. The act of transparently providing this information establishes a sense of trust between the analysts and the disclosing entities. Meanwhile, the deficiency in communicating enough information about ESG integration increases the ambiguity around the ESG performance of the entity and may significantly impact the rating and the analysts' views.

Eventually, it is worth noting that the social factors related to human capital are relevant for both clusters (best and worst), but more significantly for the "best" one. Furthermore, the other topics related to all environmental issues, some social topics ("Health and Safety" and "Consumer Protection"), and "Real Estate Green Finance" are less frequently mentioned by the analysts, as they might have been considered to play a minor role in explaining ESG risk performance. The sentiment is still in line with the ESG scores, that is positive for the best performing entities and negative for the worst ones.

The implications of this discussion can be summarized in the following points that can be of great importance for financial institutions' management with regards to their ESG performance. Based on the topic analysis, it can be suggested that governance-related aspects are the ones that attract rating analysts' attention, specifically "Business ethics", "Product governance transparency", and "Ethical conduct and mis-selling". Hence, financial institutions are called to pay great attention to foster morally accepted activities, commitment to transparent disclosure in order to instill trust and confidence in clients and stakeholders. According to the literature, these aspects not only contribute to a better ESG score, but it may also have a direct and/or an indirect impact on the institutions' market valuation. Managing the latter based on the guidelines of the ESG directives and the code of best practice is important for rating analysts as they dedicate a substantial proportion of their recommendations to these issues. The second implication derives from the second part of the present investigation. A closer look to the best and the worst performing financial institutions provides more precise indications on how the ESG best (worst)-scored financial institutions manage their ESG performance and what aspect they manage well (badly). The best ESG performing financial institutions show to be actively committed to the integration of ESG in decision making processes, the code of best practice in governance and disclosure transparency. Thus, financial players are suggested to follow the lead of best performing entities and focus on the previously mentioned aspects in terms of functional strategies and transparent disclosures. Furthermore, data privacy and security, even though classified among social topics, significantly attract analysts' attention, and should be closely monitored by financial entities, since it is a data sensitive sector. Finally, it is important to actively disclose ESG activities in a transparent way as the more information is available for rating agencies the more precise the scoring is and the better ESG commitment is reflected in analysts' views.

5. Conclusion

There is a growing body of literature discussing the importance of ESG factors and the way they are scored and evaluated. Hence, ESG disclosures, and especially textual ones, have become an interesting material worthy of deeper investigations. Multiple studies performed content analysis using both, traditional and innovative research tools, focusing on the ESG reports, news, and social media platforms to study the issue. However, to our knowledge, this paper is the first to examine ESG rating analysts' views as input data in order to highlight the main ESG features discussed by the analysts across 11 economic sectors. The main goal is to shed the light on the differences between the financial and the other economic sectors in terms of ESG features and to identify what drives the best and the worst ESG ratings.

This study performs a topic modeling and a sentiment analysis to identify the content of analysts' opinions on the companies' ESG performance and to uncover the embedded sentiment whether positive or negative associated with each ESG feature. The combination of both methodologies provides a broad framework where ESG views expressed with a particular sentiment are connected to a particular topic distribution. Furthermore, analysts' views are associated with the best and worst ESG performing financial entities, which permits the extraction of the ESG key factors that contribute to a positive evaluation.

Our findings indicate that the breadth and complexity of the whole ESG analysts' views can be summarized in 13 main topics which combine differently across economic sectors and demonstrate significant sectoral diversity in ESG concerns. The environment-related topics, such as sustainable supply chains and environmental risks, are predominantly associated with sectors like Energy, Utilities, and Materials, which have a significant environmental footprint. In the financial sector however, environmental issues have minimal weight since financial institutions have a negligible direct impact on environmental aspects. Meanwhile, they play a crucial role in determining relevant criteria for granting access to financing, which is prominently featured in governance topics related to ESG integration. Indeed, governance issues play a central role in financial institutions' evaluation due to the regulatory requirements linked to capital management, financial stability, and risk monitoring. Despite its historical presence in the business, management and finance fields, business ethics and ethical conduct are still of high priority for rating agencies as they can directly and indirectly influence not only the ESG score but also the market valuation of a subject entity.

Upon analyzing the topics emerging from the best and the worst performing financial entities in terms of ESG commitment, analysts draw the attention to three main issues that may contribute to a good/bad ESG rating. ESG integration in the decision-making processes, data privacy and transparent active disclosures. Accordingly, the implication of these findings can be summarized in the following suggestions. Financial institutions are called to strengthen their ESG commitment to foster morally accepted activities, further engage in transparent disclosure and improve data privacy mechanisms in order to instill trust and confidence in clients and stakeholders. In fact, when comparing key ESG factors relative to the best and worst ESG performance, analysts bring to attention that following the code of best practice is positively regarded and "awarded" whereas mis-selling and unethical conduct are "heavily" penalized.

The present investigation makes a three-fold contribution to the debate on ESG. First, we combine novel research instruments, topic modeling and sentiment analysis, to extract latent information embedded in ESG ratings analysts' views. Thus, this paper provides managers with deeper practical insights about what matters for ESG evaluations. Second, the integration of qualitative and quantitative inputs in a unique framework offers stakeholders a comprehensive overview on the major factors influencing ESG performance across 11 economic sectors and specifically in the financial industry. Third, we bring to the debate a new type of textual information, ESG rating analysts' opinions, that, to our knowledge, has not been studied so far.

CRediT authorship contribution statement

Marco Mandas: Conceptualization, Methodology, Data curation, Software, Validation, Writing – original draft preparation. **Oumaima Lahmar:** Conceptualization, Methodology, Software, Resources, Investigation, Writing – original draft preparation. **Luca Piras:** Conceptualization, Supervision, Validation, Investigation, Writing – review & editing, Revision. **Riccardo De Lisa:** Conceptualization, Supervision, Validation, Investigation, Writing – review & editing, Revision.

Data availability

Data will be made available on request.

Appendix A. Topic top terms concordances

TOPIC	TERM	CONCORDANCES
Sustainable Energy	carbon	operation, exposure, company's, average, issue, management, regulation, intensity, strict, risk
	energy	cost, renewable, source, company, efficient, increase, realize, short, transition, price
	emission	ghg, global, reign, wake, reduce, company, regulation, view, progress, slow
	ghg	emission, global, momentum, reign, wake, reduce, reduction, view, progress, slow
	reduce	emission, ghg, company, progress, slow, technology, energy, initiative, electricity, process
	regulation	change, climate, risk, relate, carbon, strict, view, company, emission, associate
	renewable	energy, source, transition, begin, power, increase, company, generation, integrate, network
	fuel	fossil, energy, company, source, emission, aircraft, combustion, management, fuel, low
	power	company, generation, renewable, coal, require, source, continuous, supply, cable, exchange
	gas	emission, greenhouse, natural, increase, production, attributable, ghg, programme, world, fuel
	target	reduction, deadline, ghg, emission, company, establish, reduce, set, air, change
	reduction	ghg, programme, emission, include, strong, target, risk, practice, decommission, progressive
	price	energy, change, result, company's, exposure, broad, risk, regulation, financial, momentum
	global	emission, ghg, momentum, political, reign, wake, 2, majority, address, aircraft
	efficiency	cost, operation, opportunity, saving, energy, increase, emission, company, crucial, tighten
	climate	change, regulation, risk, relate, term, company, increase, associate, food, contextual
	generation	power, technology, electricity, coal, consume, network, renewable, significantly, attributable, heat
	source	energy, renewable, transition, begin, fuel, emission, ghg, intensive, main, feedstock
	intensity	carbon, trend, affect, business, energy, extent, model, company, company's, configuration
	electricity	energy, cost, emission, ghg, consumption, purchase, process, benefit, decrease, future
Environmental and Climate Change Risk	emission	waste, effluent, exposure, air, average, issue, company's, management, excessive, community
	waste	emission, effluent, issue, company, view, produce, volume, mineral, similar, reduction
	community	relation, project, exposure, average, issue, address, view, scrutiny, company's, management
	effluent	emission, waste, issue, management, view, average, company's, exposure, medium, address

(continued on next page)

(*continued*)

TOPIC	TERM	CONCORDANCES
	environmental	management, system, certify, include, practice, impact, external, industry, standard, due
	air	emission, noise, excessive, community, trigger, vibration, water, programme, waste, volume
	project	community, delay, impact, protest, disruption, operation, shutdown, significant, plan, address
	relation	community, issue, average, company, company's, exposure, management, view, similar, moderately
	mine	produce, waste, amount, process, development, heavy, machinery, nox, site, sox
	significant	project, community, shutdown, mine, company's, exposure, impact, alter, landscape, generate
	spill	release, oil, contaminate, air, due, emission, fine, land, levy, nox
	penalty	lawsuit, regulatory, severe, company's, exposure, regulation, relevant, result, hundred, million
	release	unintended, spill, environmental, exceed, occur, regulator, pollution, accidental, crude, due
	million	hundred, dollar, cost, delay, company, develop, clean, include, penalty, compensation
	water	waste, produce, tailings, thicken, air, emission, report, company, water, source
	delay	protest, project, permit, difficulty, dollar, million, community, industry, abandon, blockade
	produce	mineral, waste, mine, dust, heavy, machinery, nox, site, sox, amount
	tailings	leak, tailings, detect, mineral, waste, average, dewatering, initiative, management, practice
	stakeholder	monitor, community, identify, system, address, asset, concern, key, life, management
	pollution	environment, generate, impact, due, financial, risk, pipeline, spill, subindustry, air
Sustainable Supply Chain	water	water, company, risk, efficiency, scarcity, clean, programme, management, comprehensive, include
	supply	chain, human, issue, exposure, management, view, low, company, disruption, risk
	resource	exposure, management, issue, average, company's, view, medium, water, risk, company
	right	human, chain, supply, issue, exposure, average, company's, management, abuse, association
	chain	supply, human, issue, company, view, low, moderately, risk, similar, implicate
	supplier	audit, engagement, ongoing, train, regular, compliance, social, standard, set, comprehensive
	material	raw, source, chain, produce, recycle, water, process, body, lotion, sunscreen
	human	supply, chain, issue, exposure, average, abuse, company's, management, association, force
	production	cost, increase, process, conflict, quantity, vast, growth, resource, water, chemical
	manufacture	component, electronic, quantity, water, product, rely, region, defect, east, require
	reduce	water, include, product, initiative, ban, develop, sale, efficiency, reduction, cost
	process	water, production, quantity, vast, manufacture, resource, company, availability, clean, experience
	chemical	hazardous, limit, substance, hazard, impact, risk, production, harmful, involve, commodity
	region	water, company, scarcity, scarce, prevalent, component, electronic, manufacturer, company's, exposure
	component	clean, electronic, component, manufacture, manufacturer, quantity, water, activity, disruption, experience
	recycle	water, reuse, drive, efficiency, reduce, company, invest, material, minimize, system
	substance	hazardous, chemical, involve, regulation, worldwide, product, personal, act, company, consequence
	electronic	component, manufacturer, issue, rely, experience, operational, prevalent, region, company, manufacture
	consumer	company, preference, additionally, expose, shift, pressure, regulator, subindustry, industry, cater
	source	exposure, issue, management, average, company's, view, medium, water, risk, company
Human Capital Management	human capital	management, issue, average, company's, company, exposure, view, inadequate, result, similar
	diversity	employee, programme, demographics, embed, improve, contribute, discrimination, limit, inefficiency, initiative
	manage	average, capital, human, company's, issue, exposure, view, medium, effectively, employee
	initiative	employee, programme, recruitment, retention, implement, diversity, roll, train, tailor, company
	support	drive, growth, innovation, strategic, workforce, depend, array, complex, product, retain
	workforce	growth, strategic, support, depend, retain, reits, disproportionately, diversify, industry, technology
	disclose	employee, demographics, information, involve, practice, average, diversity, alike, company, train
	growth	support, drive, innovation, strategic, workforce, foster, capital, deliver, human, inability
	recruitment	initiative, programme, retention, employee, diversity, embed, tailor, workforce, option, ownership
	depend	support, workforce, reits, company, design, estate, real, subindustry, industry, bank
	client	investor, view, alike, positively, reits, company's, exposure, company, diversity, improve
	skill	deficit, highly, require, engineer, software, employee, compliance, failure, incentivize, develop
	involve	demographics, disclose, employee, information, practice, average, subindustry, construction, effectively, incentivize
	innovation	support, drive, growth, strategic, foster, trust, loyalty, array, complex, client
	legal	compliance, programme, view, employee, company, diversity, improve
	drive	support, growth, innovation, strategic, foster, trust, complex, client, loyalty,
	talent	acquire, integrate, pose, risk, material, newly, respect, recruitment, board, human
	strategic	drive, growth, innovation, support, workforce, retain, strong
	improve	diversity, discrimination, limit, failure, inefficiency, operational, average, subindustry, initiative, access

(*continued on next page*)

(*continued*)

TOPIC	TERM	CONCORDANCES
Human Capital Empowerment	human capital	average, issue, exposure, management, company's, view, medium, strong, company
	labour	experience, company, relation, strain, require, represent, process, manufacture, skill, low
	development	programme, human, capital, strong, company, include, talent, initiative, engagement, opportunity
	initiative	mentorship, wide, widespread, talent, diversity, development, flexible, recruitment, workforce, labour
	turnover	employee, impact, low, rate, form, line, bottom, motivate, motivation, turnover
	train	company, employee, labour, relation, include, activity, hr, experience, process, unmotivated
	retain	qualify, attract, recruit, employee, fail, innovation, stagnate, personnel, sound, strategy
	talent	include, initiative, development, average, practice, recruitment, attraction, plan, stock, talent
	diversity	initiative, programme, workforce, labour, train, basic, flexible, standard, monitor, employee
	highly	company, industry, seek, practitioner, product, skill, expose, personnel, issue, capital
	rate	average, capital, issue, subindustry, turnover, medium, employee, company, impact, recruit
	develop	human, capital, programme, strong, company, talent, employee, include, initiative, recruit
	experience	labour, company, relation, customer, diminish, insurance, manufacture, process, significantly, strain
	attract	retain, include, qualify, company, fail, employee, talent, option, plan, practice
	staff	production, facility, staff, include, cost, company, compliance, market, professional, sale
	skill	highly, require, industry, labour, low, design, manufacture, personnel, company, machinery
	wage	hour, relate, strike, freedom, company, experience, insurance, restructure, fair, code
	discrimination	policy, association, freedom, level, minimum, salary, strong, labour, represent, workplace
	personnel	industry, highly, skill, design, product, retain, key, airline, attendant, mechanic
Data Privacy and Security	datum	security, privacy, breach, exposure, issue, management, average, company's, datum, company
	security	datum, privacy, company, exposure, issue, view, management, medium, risk, similar
	privacy	datum, security, management, issue, company's, exposure, view, average, medium, risk
	information	sensitive, identifiable, personally, datum, volume, security, highly, customer, social, ensure
	breach	datum, increase, frequency, company, disruption, service, cost, total, loss,
	process	company, volume, store, identifiable, personally, receive, security, define, ensure, information
	cybersecurity	exposure, due, regulate, sector, event, market, negative, react, share, typically
	sensitive	datum, information, identifiable, personally, customer, highly, market, react, access, industry
	financial	industry, datum, 2010, 3, experience, sector, record, credit, drive, history
	cyberattacks	datum, exposure, frequency, breach, prime, security, target, collect, risk, industry
	assessment	privacy, risk, security, infrastructure, technology, update, view, regular, conduct, base
	personal	identifiable, information, sensitive, process, volume, datum, intellectual, property, decision, regard
	threat	datum, remove, system, damage, relate, repair, emerge, exist, plan, address
	collect	datum, customer, company, behaviour, process, telecom, transmit, cyberattacks, store, complement
	sector	cybersecurity, due, regulate, 2010, 3, experience, financial,
	technology	assessment, evolve, infrastructure, risk, update, company, proactive, hardware, issue, product
	identifiable	personally, information, datum, process, sensitive, volume, property, intellectual, addition, collect
	due	cybersecurity, exposure, regulate, sector, breach, customer, datum, dollar, million, ten
	regulate	cybersecurity, due, sector, confidential, information, manipulation, approach, base, datum, insurance
	level	management, oversight, environment, system, datum, issue, security, strong, assign, managerial
Consumer Protection	quality	safety, management, product, machine, regulation, system, standard, include, practice, determine
	governance	product, issue, average, company, company's, exposure, management, view, similar, medium
	safety	quality, product, machine, include, management, system, affect, directly, certification, implement
	standard	management, quality, expensive, extensive, lead, recall, regulation, international, iso, practice
	recall	product, lead, customer, complaint, expensive, increase, regulatory, scrutiny, standard, 9001
	test	quality, 17025, adverse, asw, certify, control, division, iso, pass, reaction
	ensure	company, quality, advance, machinery, stage, alignment, regulatory, standard, manufacture, product
	certification	management, system, safety, 9001, iso, facility, accord, international, programme, ethical
	food	administration, authority, drug, fda, safety, concern, lead, producer, quality, issue
	iso	9001, certification, proactive, product, recall, system, certify, 9000, additionally, company
	failure	adhere, deceptive, extensive, false, market, regulation, meet, regulatory, requirement, addition
	trial	clinical, datum, quality, machine, machinery, 3, complaint, customer, handle, transparency
	sale	ethical, incentive, market, personnel, regular, structure, train, company's, exposure, product
	lawsuit	customer, company's, escalate, frequently, impact, affect, compliance, cost, scrutiny, action
	implement	include, management, practice, quality, safety, system
	clinical	datum, trial, complaint, customer, handle, transparency, regard, transparent, company
	manufacture	effect, include, industry, irregularity, unanticipated, pharmaceutical, biotech, company, gmp, practice
	defect	threaten, customer, industrial, injury, result, safety, machine, purchase, machinery, product
	conduct	assessment, risk, periodic, product, weak, programme, evidence, food, safety,

(*continued on next page*)

(*continued*)

TOPIC	TERM	CONCORDANCES
Health and Safety	safety	health, occupational, management, issue, company, exposure, view, include, practice, record
	health	safety, occupational, management, average, issue, company's, exposure, view, medium, train
	occupational	health, safety, exposure, average, issue, company's, management, disease, injury, illness
	fatality	employee, limit, result, workplace, fatality, injury, morale, productivity, prone, commonplace
	incident	safety, investigation, temporarily, occupational, diesel, fuel, rate, lose, time, company
	contractor	train, safety, employee, oversight, tend, fatality, lead, negative, view, monitor
	train	safety, health, contractor, extend, oversight, employee, regular, include, management, system
	accident	avoid, essential, integrity, maintain, productivity, workforce, damage, main, ship, financial
	performance	safety, poor, correlate, environment, financial, industrial, labour, record, strike, trigger
	injury	disease, financial, impact, occupational, illness, fatality, result, typically, industry, company
	rate	safety, issue, average, subindustry, company, medium, key, environment, hazardous, employee
	operational	inefficiency, result, operation, tend, fatality, avoid, essential, integrity, response, preparedness
	worker	compensation, family, lawsuit, expose, ability, attract, exposure, company's, adequate, equipment
	gas	oil, nature, volatile, issue, key, petrochemical, company, natural, safety, contractor
	emergency	include, management, preparedness, risk, response, train, check, medical, safety, system
	standard	safety, include, 18001, ohsas, health, maintain, alignment, management, certify, compliance
	regular	employee, train, safety, check, inspection, upgrade, assessment, equipment, risk, protective
	operate	safety, company, key, environment, hazardous, offshore, rig, employee, procedure, assessment
	fine	result, fatality, typically, lawsuit, low, morale, operation, shutdown, penalty, company
	hazardous	remote, environment, company, operate, chemical, contact, diesel, condition, include, perform
Business Ethich and Code of Conduct	ethic	business, issue, infraction, company's, exposure, company, average, management, view, include
	corruption	bribery, risk, company, management, expose, company's, exposure, issue, view, average
	bribery	corruption, risk, company's, exposure, expose, average, issue, industry, management, medium
	ethical	include, infraction, broad, market, country, operate, insider, practice, trade, commitment
	whistleblower	report, confidentially, treat, programme, average, subindustry, company's, unclear, anonymous, permit
	bank	service, trustworthy, diversify, play, vital, fine, receive, record, claw, discipline
	fine	ethic, infraction, receive, record, bank, investment, lawsuit, expose, hit, million
	tax	evasion, compliance, launder, money, regulation, diligence, due, practice, account, competitive
	launder	money, sanction, violation, regulatory, regulation, ethic, international, link, programme, compliance
	account	country, fund, movement, multiple, fraud, practice, relate, taxation, tax, anti
	infraction	ethical, include, ethic, fine, receive, record, broad, manipulation, market, bonus
	investigation	corruption, award, contract, expose, anti, lawsuit, practice, relate, abuse, company
	range	service, product, offer, complex, financial, bank, diversify, relate, business, ethic
	money	launder, sanction, violation, regulation, ethic, international, link, programme, compliance, promote
	country	operate, ethical, include, infraction, multiple, account, casino, fund, vulnerable, develop
	intellectual	property, anti, competitive, relate, company, practice, issue, account, business, emerge
	financial	complex, offer, product, range, service, effect, knock, performance, company, guilty
	conflict	issue, insider, trade, expose, highly, surround, launder, money, business, ethic
Product Governance Transparency	disclose	company, price, list, publicly, policy, change, average, programme, employee, product
	conduct	risk, customer, framework, incorporate, insurer, train, code, company, impact, social
	major	controversy, implicate, issue, relate, access, barrier, company, disease, drug, federal
	commitment	customer, profitability, compensate, insure, tension, honour, honor, include, insurer, pay
	access	service, basic, issue, average, exposure, management, company's, ensure, medicine, company
	potential	risk, company, price, scrutinize, 19, covid, exposure, legal, novartis, regulatory
	global	santander, digital, platform, support, transformation, customer, product, access, addition, bank
	train	market, increase, budget, care, develop, health, public, system, employee, conduct
	price	company, pressure, increase, drug, product, adjust, provide, biotech, coverage, regulatory
	lack	disclosure, satisfaction, clarity, inform, measure, regulatory, risk, zurich, company, product
	usd	bn, usd, 1, 100, 300, amount, billion, programme, relate, settlement
	initiative	improve, company, access, average, medicine, icbc, quality, service, disclose, include
	strategy	product, pursue, aggressive, financial, gain, compliance, explore, governance, regulatory, consumer
	significant	company's, exposure, fine, result, settlement, company, price, average, risk, issue
	base	company, customer, price, product, expand, performance, engage, pharmaceutical, care, health
	million	result, hundred, penalty, company's, life, customer, eur, million, company, fine
	china	customer, ensure, liberalization, protection, rate, share, market, access, china, drug
	provide	care, health, capacity, product, price, local, performance, agreement, build, public
	annual	complaint, report, market, offering, increase, 6, average, h1, resolution, semi
	committee	compliance, ethic, price, probe, bank, complaint, customer, dedicate, drug, establish
ESG integration in investment and financing decision	integration	esg, financials, issue, average, company's, exposure, management, view, medium, address
	investment	product, decision, esg, integrate, align, deliver, demand, depend, innovative, moral
	financials	integration, esg, issue, company, exposure, management, view, medium, similar, address
	bank	diversify, product, array, comprehensive, criticism, lend, level, offer, operation, subject
	responsible	finance, opportunity, asset, bank, relate, risk, expectation, increasingly, public, shift

(*continued on next page*)

(*continued*)

TOPIC	TERM	CONCORDANCES
	opportunity	responsible, business, capture, array, relate, risk, multifaceted, asset, manager, finance
	asset	manager, management, opportunity, asset, increase, increasingly, responsible, company's, exposure, source
	project	energy, lend, controversial, industry, efficiency, renewable, coal, mine, likewise, company's
	diversify	bank, product, array, comprehensive, criticism, lend, offer, operation, subject, multifaceted
	finance	responsible, bank, opportunity, relate, expectation, increasingly, public, suggest
	invest	energy, sustainability, esg, company, bond, green, renewable, cater, consumer, mind
	offer	product, bank, comprehensive, diversify, deliver, depend, innovative, investment, subindustry, client
	social	environmental, challenge, tackle, business, equity, fund, asset, esg, governance, integration
	environmental	social, challenge, tackle, business, asset, manager, financials, governance, integration, involve
	increasingly	asset, manager, finance, responsible, suggest, investor, millennial, term, client, offer
	lend	project, bank, controversial, encounter, footprint, geographical, industry, operation, risk, diversify
	controversial	project, encounter, industry, lend, risk, dam, hydroelectric, manufacturer, mine, palm
	sustainable	carbon, estate, investment, low, portfolio, real, financial, align, business, development
	financial	integration, esg, issue, company, exposure, management, view, medium, similar, address
	Integrate	esg, decision, investment, customer, demand, evolve, bank, consideration, reduce, efficiency
Real estate green finance	build	green, certification, development, efficient, energy, system, premium, space, provide, safety
	property	risk, change, equip, ill, invest, affect, sale, develop, adversely, aid
	revenue	company's, exposure, impact, reit, sale, adversely, affect, capital, company, devaluation
	real estate	erode, finance, restriction, result, tight, esg, income, integrate, manager
	tenant	cost, increase, productivity, satisfaction, utility, worker, safety, build, poor, ability
	green	build, certification, development, efficient, energy, system, building, premium, space, provide
	safety	tenant, quality, company, build, poor, compliance, legislative, requirement, building, hurt
	portfolio	affect, asset, change, climate, estate, negative, performance, real, strand, dividend
	term	management, factor, investment, process, risk, income, manager, owner, reits, climate
	climate	change, estate, risk, additionally, regulation, invest, portfolio, result, tight, unitholders
	building	green, premium, space, provide, damage, likewise, vulnerability, unitholders, company, quality
	development	build, certification, comprehensive, green, robust, aid, assessment, company, decision, property
	integration	esg, financials, issue, average, company's, exposure, management, view, medium, address
	financials	esg, integration, issue, company, exposure, management, view, medium, moderately, similar
	standard	assessment, esg, integrate, risk, estate, real, company, aid, build, construction
	factor	esg, include, platform, risk, term, decision, key, tenant, client
	change	climate, property, risk, invest, portfolio, develop, estate
	affect	asset, negative, performance, portfolio, strand, lower, potentially, property, adversely, capital
	ensure	annually, audit, compliance, legislative, property, quality, include, initiative, practice, product
Ethics Conduct and Misseling	governance	product, management, issue, company, average, company's, exposure, view, refer, similar
	quality	management, control, internalize, product, service, financial, thousand, practice, system, branch
	financial	product, management, offer, quality, service, company, compare, impact, low, subindustries
	client	interaction, individual, thousand, branch, view, company, control, cross, transaction, favour
	bank	management, incentivizes, institute, oversight, practice, regulatory, governance, product, quality, refer
	offer	financial, product, service, include, misconduct, type
	lend	false, market, discriminatory, practice, predatory, misconduct, include, type, investor, mislead
	control	internalize, quality, thousand, management, system, branch, client, advisor, individual, interaction
	reputational	demonstrate, operational, risk, bank, company's, exposure, loss, mismanagement, company, controversy
	poor	disclosure, foreclosure, illegal, investor, lend, mislead, practice, consequence, crisis, demonstrate
	disclosure	poor, foreclosure, illegal, investor, mislead, practice, excessive, include, product, investment
	oversight	average, institute, regulatory, subindustry, incentivizes, bank, asset, manager, company, financial
	type	include, misconduct, market, discriminatory, lend, offer, service, false, illegal, model
	misconduct	include, type, market, lend, discriminatory, practice, offer, poor, false, illegal
	fine	mismanagement, possibility, risk, expensive, penalty, company's
	false	lend, include, market, discriminatory, practice, predatory, disclosure, misconduct, poor, type
	refer	governance, product, management, bank, asset, managers, company, regional, requirement, service
	individual	client, interaction, thousand, branch, company, control, quality
	interaction	client, individual, branch, view, company, thousand, average, control

This table represents the cluster of word concordances for the top terms in each topic.

Appendix B. Business model characteristics and the corresponding ESG feature for each activity sector

Key Business Model Factors				
Sectors	**Key Activities**	**Main revenue.**	**Main Costs**	**ESG key topic***
Financial	Financial services, insurance, asset management, investment banking.	Interest from loans, services' fees, commissions, trading activities, investment returns.	Operating expenses, employee salaries, infrastructure, regulatory compliance, credit provisions.	9, 10, 11, 13
Real estate	Developing, owning, operating, and leasing properties.	Rental incomes, real estate sales and transactions, property management fees, development profits.	Property maintenance and improvement, taxes, mortgage interest expenses, and marketing.	12
TLC Services	Communication services, voice, data, internet, and mobile services.	Service subscriptions, voice and data usage fees, equipment sales, roaming charges, other services.	Infrastructure, spectrum licenses, customer acquisition, customer support, maintenance.	6 and 10
Health care	Providing medical services, healthcare products, and pharmaceuticals.	Healthcare service fees, sales of pharmaceuticals and medical devices, health insurance premiums, research grants.	Medical staff salaries, research and development, pharmaceuticals and medical supplies, regulatory compliance, facility maintenance.	8
IT	Developing, implementing, and managing information technology solutions and services.	Software sales and licensing, IT consulting and professional services, cloud computing services, hardware sales.	R&D, employee salaries and benefits, data centers and infrastructure, marketing.	4, 6 and 9
Industrials	Manufacturing and producing industrial goods and services.	Sales of machinery, equipment, aerospace products, engineering services, construction, industrial technology.	Manufacturing, labor, R&D, marketing and distribution, regulatory compliance.	1, 5, and 9
Consumer Discretionary	Manufacturing and selling goods and non-essential services.	Sales of automobiles, apparel, household goods, recreational products, leisure services, entertainment.	Cost of goods sold, marketing, R&D.	3 and 5
Consumer Staples	Manufacturing and selling essential household consumer products.	Sales of food, beverages, personal care products, household essentials, and other everyday consumer goods.	Raw materials, production, distribution and logistics, marketing.	1, 3 and 7
Utilities	Generating, transmitting, and distributing essential public services.	Sales of electricity, gas, water, and related services.	Energy production, maintenance and operation of infrastructure, regulatory compliance, and customer service.	1, 2 and 8
Energy	Exploration, extraction, production, refining, and distribution of energy.	Sales of energy resources, natural gas, coal, and renewable energy sources, electricity generation, energy trading, and energy services.	Exploration and production, refining and processing, transportation and logistics, maintenance and operation of energy infrastructure, regulatory compliance, environmental and safety measures.	1,2, 8
Materials	Extracting, processing, and manufacturing raw materials and basic resources.	Sales of metals, minerals, chemicals, construction materials, and other raw or processed materials.	Raw materials, mining and extraction, processing, transportation, environmental compliance.	1,2, and 3

*ESG issues extracted by the topic modeling corresponding to the activity sector.

This table represents the sectors of interest, the main business model characteristics including key activities to generate revenues and support costs and the corresponding ESG issue that may be influenced. The corresponding ESG features are the topics extracted by the topic modeling.

References

Amel-Zadeh, A., Serafeim, G., 2018. Why and how investors use ESG information: evidence from a global survey. Financ. Anal. J.

Amin, M.H., Mohamed, E.K.A., Elragal, A., 2021. CSR disclosure on Twitter: Evidence from the UK. Int. J. Account. Inf. Syst. 40, 100500 https://doi.org/10.1016/j.accinf.2021.100500.

Asmussen, C.B., Møller, C., 2019. Smart literature review: a practical topic modelling approach to exploratory literature review. J. Big Data 6, 93. https://doi.org/10.1186/s40537-019-0255-7.

Aureli, S., Gigli, S., Medei, R., Supino, E., 2020. The value relevance of environmental, social, and governance disclosure: Evidence from Dow Jones Sustainability World Index listed companies. Corp. Soc. Responsib. Environ. Manag. 27, 43–52. https://doi.org/10.1002/csr.1772.

Baier, P., Berninger, M., Kiesel, F., 2020. Environmental, social and governance reporting in annual reports: A textual analysis. Financ. Mark. Inst. Instrum. 29, 93–118. https://doi.org/10.1111/fmii.12132.

Bapat, S.R., Kothari, S., Bansal, R., 2022. Sentim. Anal. ESG Discl. Stock Mark. https://doi.org/10.48550/arXiv.2210.00731.

Barberis, N., Shleifer, A., Vishny, R., 1998. A model of investor sentiment. J. Financ. Econ. 37.

Benvenuto, M., Aufiero, C., Viola, C., 2023. A systematic literature review on the determinants of sustainability reporting systems. Heliyon 9, e14893. https://doi.org/10.1016/j.heliyon.2023.e14893.

Bhimavarapu, V.M., Rastogi, S., Abraham, R., 2022. The influence of transparency and disclosure on the valuation of Banks in India: the moderating effect of environmental, social, and governance variables, shareholder activism, and market power. J. Risk Financ. Manag. 15, 612. https://doi.org/10.3390/jrfm15120612.

Blei, D.M., Ng, A.Y., Jordan, M.I., 2003. Latent dirichlet allocation. J. Mach. Learn. Res. 3, 993–1022.

Bradshaw, M., Ertimur, Y., O'Brien, P., 2017. Financial analysts and their contribution to well-functioning capital markets. Found. Trends® Account. 11, 119–191. https://doi.org/10.1561/1400000042.

Bradshaw, M.T., 2009. Analyst information processing, financial regulation, and academic research. Account. Rev. 84, 1073–1083.

Brooks, C., Oikonomou, I., 2018. The effects of environmental, social and governance disclosures and performance on firm value: A review of the literature in accounting and finance. Br. Account. Rev., Eff. Environ., Soc. Gov. Discl. Perform. Firm Value 50, 1–15. https://doi.org/10.1016/j.bar.2017.11.005.

Chen, Y., Li, T., Zeng, Q., Zhu, B., 2023. Effect of ESG performance on the cost of equity capital: Evidence from China. Int. Rev. Econ. Financ. 83, 348–364. https://doi.org/10.1016/j.iref.2022.09.001.

Deegan, C., Rankin, M., Tobin, J., 2002. An examination of the corporate social and environmental disclosures of BHP from 1983–1997: A test of legitimacy theory. Account. Audit. Account. J. 15, 312–343. https://doi.org/10.1108/09513570210435861.

Forssbaeck, J., Oxelheim, L., 2014. The Oxford Handbook of Economic and Institutional Transparency. Oxford University Press,.

Giles, O., Murphy, D., 2016. SLAPPed: the relationship between SLAPP suits and changed ESG reporting by firms. Sustain. Account. Manag. Policy J. 7, 44–79. https://doi.org/10.1108/SAMPJ-12-2014-0084.

Gillan, S.L., Koch, A., Starks, L.T., 2021. Firms and social responsibility: A review of ESG and CSR research in corporate finance. J. Corp. Financ. 66, 101889 https://doi.org/10.1016/j.jcorpfin.2021.101889.

Goloshchapova, I., Poon, S.-H., Pritchard, M., Reed, P., 2019. Corporate social responsibility reports: topic analysis and big data approach. Eur. J. Financ. 25, 1637–1654. https://doi.org/10.1080/1351847X.2019.1572637.

de Haan, J., Vlahu, R., 2016. Corporate Governance of Banks: a survey. J. Econ. Surv. 30, 228–277. https://doi.org/10.1111/joes.12101.

Hahn, R., Kühnen, M., 2013. Determinants of sustainability reporting: a review of results, trends, theory, and opportunities in an expanding field of research. J. Clean. Prod. 59, 5–21. https://doi.org/10.1016/j.jclepro.2013.07.005.

He, R., Chen, X., Chen, C., Zhai, J., Cui, L., 2021. Environmental, Social, and Governance Incidents and Bank Loan Contracts. Sustainability 13, 1885. https://doi.org/10.3390/su13041885.

Hu, M., Liu, B., 2004. Mining and summarizing customer reviews. In: Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '04. Association for Computing Machinery, New York, NY, USA, pp. 168–177. https://doi.org/10.1145/1014052.1014073.

Jackson, G., Bartosch, J., Avetisyan, E., Kinderman, D., Knudsen, J.S., 2020. Mandatory Non-financial Disclosure and Its Influence on CSR: An International Comparison. J. Bus. Ethics 162, 323–342. https://doi.org/10.1007/s10551-019-04200-0.

Jayarathna, C.P., Agdas, D., Dawes, L., Miska, M., 2021. Exploring sector-specific sustainability indicators: a content analysis of sustainability reports in the logistics sector (ahead-of-print). Eur. Bus. Rev.. https://doi.org/10.1108/EBR-02-2021-0047.

Jenkins, H., Yakovleva, N., 2006. Corporate social responsibility in the mining industry: Exploring trends in social and environmental disclosure. J. Clean. Prod. 14, 271–284. https://doi.org/10.1016/j.jclepro.2004.10.004.

Kiesel, F., Lücke, F., 2019. ESG in credit ratings and the impact on financial markets. Financ. Mark. Inst. Instrum. 28, 263–290. https://doi.org/10.1111/fmii.12114.

Lang, M.H., Lundholm, R.J., 1996. Corporate disclosure policy and analyst behavior. Account. Rev. 71, 467–492.

Lee, H., Lee, S.H., Lee, K.R., Kim, J.H., 2023. ESG discourse analysis through BERTopic: comparing news articles and academic papers. Comput. Mater. Contin. 75, 6023–6037. https://doi.org/10.32604/cmc.2023.039104.

Li, F., 2008. Annual report readability, current earnings, and earnings persistence. J. Account. Econ. 45, 221–247. https://doi.org/10.1016/j.jacceco.2008.02.003.

Lin, C., He, Y., 2009, Joint sentiment/topic model for sentiment analysis.

Lindgren, C., Huq, A.M., Carling, K., 2021. Who are the intended users of CSR reports? Insights from a data-driven approach. Sustainability 13, 1070. https://doi.org/10.3390/su13031070.

Liu, B., 2020. Sentiment Analysis: Mining Opinions, Sentiments, and Emotions. Cambridge University Press,.

Liu, M., Luo, X., Lu, W.-Z., 2023. Public perceptions of environmental, social, and governance (ESG) based on social media data: Evidence from China. J. Clean. Prod. 387, 135840 https://doi.org/10.1016/j.jclepro.2022.135840.

Loughran, T., McDonald, B., 2020. Textual analysis in finance. Annu. Rev. Financ. Econ.

Madison, N., Schiehll, E., 2021. The effect of financial materiality on ESG performance assessment. Sustainability 13, 3652. https://doi.org/10.3390/su13073652.

Ning, X., Yim, D., Khuntia, J., 2021. Online sustainability reporting and firm performance: lessons learned from text mining. Sustainability 13, 1069. https://doi.org/10.3390/su13031069.

Park, J., Choi, W., Jung, S.-U., 2022. Exploring trends in environmental, social, and governance themes and their sentimental value over time. Front. Psychol. 13, 890435 https://doi.org/10.3389/fpsyg.2022.890435.

Riedl, A., Smeets, P., 2017. Why do investors hold socially responsible mutual funds? J. Financ. 72, 2505–2550. https://doi.org/10.1111/jofi.12547.

Rjiba, H., Saadi, S., Boubaker, S., Ding, X. (Sara), 2021. Annual report readability and the cost of equity capital. J. Corp. Financ. 67, 101902 https://doi.org/10.1016/j.jcorpfin.2021.101902.

Rouen, E., Sachdeva, K., Yoon, A., 2022. The Evolution of ESG Reports and the Role of Voluntary Standards. SSRN Electron. J. https://doi.org/10.2139/ssrn.4227934.

Russell, B., 1959. The Realism of Bertrand Russell. Aryan Path.

Seo, H., Jo, D.H., Pan, Z., 2022. Big Data Analysis of 'ESG' News Using Topic Modeling. 2022 IEEE/ACIS 7th International Conference on Big Data, Cloud Computing, and Data Science (BCD). Presented at the 2022 IEEE/ACIS 7th International Conference on Big Data. Cloud Computing, and Data Science (BCD),, pp. 183–187. https://doi.org/10.1109/BCD54882.2022.9900604.

Serafeim, G., Yoon, A., 2022. Which corporate ESG news does the market react to? Financ. Anal. J. 78, 59–78. https://doi.org/10.1080/0015198X.2021.1973879.

Shen, C., Li, P., 2023. Green housing on social media in China: A text mining analysis. Build. Environ. 237, 110338 https://doi.org/10.1016/j.buildenv.2023.110338.

Székely, N., vom Brocke, J., 2017. What can we learn from corporate sustainability reporting? Deriving propositions for research and practice from over 9,500 corporate sustainability reports published between 1999 and 2015 using topic modelling technique. PLOS ONE 12, e0174807. https://doi.org/10.1371/journal.pone.0174807.

Teece, D.J., 2010. Business Models, Business Strategy and Innovation. Long. Range Plann., Bus. Models 43, 172–194. https://doi.org/10.1016/j.lrp.2009.07.003.

Zhao, W., Chen, J.J., Perkins, R., Liu, Z., Ge, W., Ding, Y., Zou, W., 2015. A heuristic approach to determine an appropriate number of topics in topic modeling. BMC Bioinforma. 16, S8. https://doi.org/10.1186/1471-2105-16-S13-S8.