The ESG Score and the financial statement: the European case

Rita Laura D'Ecclesia[*] Susanna Levantesi[†] Kevyn Stefanelli[‡]

First Version: Dec 20, 2022;
This Version: July 31, 2023

Abstract

Environmental, Social, and Governance (ESG) issues play a central role in the foundation of business strategies; for a set of managers and financial institutions, ESG is now an essential part of risk management. Measuring the ESG commitment of companies using ESG ratings has been questioned but is becoming common for stock selections and asset management strategies. In this paper, we want to analyze the role played by ESG ratings in the performance of European-listed companies. We use ESG ratings and corporate balance sheet information for the European companies constituent the various stock indices, to identify which are the company structural drivers explaining the ESG score for the European companies. We use a Machine Learning approach to take into account non-linear relationships and to determine, for each country, the main drivers of the ESG score, providing insights for sustainable investment selections. We also study how the Environment, Social and Governance components affect the various economic sectors and explain their recent dynamics in Europe.

Keywords: ESG Score, Machine Learning, Corporate Finance, Sustainability.

1 Introduction

Western economies are dealing with a period of high uncertainty caused by increasing inflation, energy market turmoil, rising interest rates, pandemic fatigue, and political uncertainty. In this context, Environmental, Social, and Governance (ESG) issues have become the current mantra: we see national and international regulators upping the ante on everything from greenwashed fund

[*]Department of Statistics, Sapienza University of Rome, Rome, Italy. E-mail: rita.decclesia@uniroma1.it
[†]Corresponding author. Department of Statistics, Sapienza University of Rome, Rome, Italy. E-mail: susanna.levantesi@uniroma1.it
[‡]Department of Statistics, Sapienza University of Rome, Rome, Italy. E-mail: kevyn.stefanelli@uniroma1.it

names to stricter climate target disclosures. ESG investing has taken center stage for asset managers and portfolio selection strategies. Modern investors are reevaluating traditional investment approaches and incorporating new risk factors that may not have been previously considered, such as flood risk and sea level rise, privacy and data security, demographic shifts, and regulatory pressures. Also investor preferences are changing: millennials - as well as women - are demanding more from their investments. They are increasingly committed to sustainability and making more environmentally conscious choices compared to previous generations (Morgan Stanley, 2017). While the intentions of investors are clear, incorporating sustainability in investment management remains challenging. In addition to traditional expected return and volatility, there is a need to include an additional measure to assess the quality of investment portfolios and their risks linked to the ESG component. The ESG performance of a firm involves multiple dimensions. Environmental aspects concern emissions and waste management, social performance encompasses workforce diversity and community involvement, while governance measures include stakeholder engagement and share voting rights. In the last two decades, researchers and financial agents have strived to uniformly measure all these aspects to identify a single number that provides information on the company's sustainable commitment. The ESG score probably represents the most successful attempt in this sense.

Overall, ESG data and metrics are still considered insufficient and often untrustworthy. They are not specifically designed to provide disclosure on the PAIs, despite often being used for this purpose. ESG ratings serve as a metric in the quest for responsible KPI. Numerous agencies have come into play in the last decade to provide company-based ratings. These are the result of a time-intensive and challenging process task for multiple renowned ESG data providers and rating agencies. This is mainly due to regulations and reporting requirements that vary across different countries and regions around the world. Rating agencies exploit a wide set of ESG parameters, or factors, which include company self-disclosures, media reports, controversies, investor risk, and more. Nowadays, ESG ratings are issued by six major agencies that boast an extensive global coverage of public and private companies: Bloomberg, Dow Jones, Institutional Shareholder Services, MSCI, Refinitiv, and Sustainalytics (Berg et al., 2022). Refinitiv was established in 2002, and claims to provide one of the most thorough ESG datasets, encompassing more than 80% of the worldwide market value over more than 630 different ESG metrics. The Refinitiv ESG score ranges from 0 to 100, where a score closer to 100 indicates a company that is most aligned with ESG principles. They include a controversy score, and a Combined Score. The controversy score is generated based on controversial ESG news/events related to the specific company happening across the globe (Refinitiv, 2022a), The combined score is calculated as the average of the ESG score and ESG controversies score when there are controversies during the fiscal year. When the controversies score is greater than ESG score, then ESG score is equal to ESG combined score.

Regulators' requirements and government policies are affecting the investment environment as well as the challenges and opportunities companies face. Regulation is influencing business as usual in the EU and is increasing in the U.S. and the Asia-Pacific (APAC) markets: from requirements for financial institutions to conduct climate stress tests, to investors getting ahead of mandatory requirements to report on the Principle Adverse Impact (PAIs) indicators considered in the Sustainable Finance Disclosure Regulation (SFDR) (European Commission, 2023b). In Europe and the UK the SFDR and the Corporate Sustainability Reporting Directive (CSRD) (European Commission, 2023a) are leading the way. These reports introduce taxonomies which help investors to discriminate the actual sustainable investment. Other countries, such as the U.S., Australia, and South Africa have all trailed behind Europe on ESG issues. In 2024, financial-market participants subject to the SFDR are subject to provide a report on the PAIs indicators associated with their portfolio holdings and monitor their changes over time. A sustainable investment, as defined by SFDR Article 2, should not cause distress under any of the mandatory PAIs. For instance, PAIs can be used to provide information regarding the sustainability of an investment when dealing with investors' preferences under the Markets in Financial Instruments Directive (MiFID II) rules.

Institutional investors are tasked to develop portfolios that withstand market volatility, respond to a growing suite of regulatory requirements, and identify sustainable growth opportunities that will deliver long-term returns to achieve the 17 United Nations (UN) Sustainable Development Goals (SDG) (United Nations, 2012). At the same time, the literature of ESG factors has grown apace. The evolving biodiversity, the variation in the social dynamics, and the development of the supply chain related risks and opportunities are broadening the understanding of what the multi-trillion-dollar sustainable finance (SF) industry can offer to investors. For most investors, ESG features are an essential part of risk management. A large part of them invests capitals in ESG-themed funds betting on the long-term growth of the sustainable business[1].

Europe remains the key driver behind the SF, where sustainability objectives and the respective legislative interventions are shaping the future of the fund industry. The 83% of global sustainable funds' net assets are located in Europe (Morningstar (2020) and Zeb (2022), powered by the Association of the Luxembourg Fund Industry (ALFI)). The net assets invested in sustainable funds, based on Morningstar's strict definition of sustainability, have reached almost EUR 2 trillion at the end of 2021, growing up by 71% from 2020. Sustainable funds represent 16% of total net assets of

[1] ESG-themed funds are defined as any fund that employs any ESG considerations in its security-selection process (Robeco, 2023).

funds domiciled in Europe, ahead of the U.S. and Asia, with only 1% and 5%, respectively. Luxembourg maintains its market leader position, with about a third of the assets managed by sustainable funds in Europe being domiciled there. Equity remains the most important asset class, making up 64% of the sustainable funds compared to 48% in conventional funds, allowing asset managers to exert a great influence on the ESG efforts of companies. A different trend is observed in the U.S.. At the beginning of 2022, USD 7.6 trillion were held by a list of 497 institutional investors, 349 money managers, and 1359 community investment institutions that include ESG criteria in their investment decision-making process. In addition, 154 institutional investors and 70 money managers controlling USD 3.0 trillion in assets under management led or co-led shareholder resolutions on ESG issues from 2020 through the first half of 2022. producing USD 8.4 trillion (representing 12% of the $66 trillion in total U.S. AUM) in sustainable investing strategies, while in 2020 they represented 33%.

As the demand for responsible investments continues to grow, investors and stakeholders need to understand the relationship between companies' performance and ESG rating, so that asset managers may use ESG ratings when building their portfolios. The aim of the paper is to identify the main drivers which explain the ESG scores assigned to European public companies. We examine economic sector distributions by country with the aim to explain ESG discrepancies. We use a composite dataset including the ESG ratings and financial statement ratios for all the companies that constitute the principal European indexes in 2011-2022. We adopt a Machine Learning (ML) approach to detect the main drivers of the ESG ratings and use interpretation tools as the feature importance, the Partial Dependent Plot, and features interaction to get a deeper understanding of the findings and then enhance the interpretability of the model.

ESG ratings show a different probability distribution in each country, which can be explained by the different economic sector composition of each index. For the UK, Italy, Germany, the Netherlands and France, the company's size drives the ESG rating. Large companies have more resources to devote funds to change the business model toward a more sustainable pattern, translating into higher ESG ratings. The second driver for these countries, with the only exception of France, is represented by Carbon Intensity; as the impacts of climate change become more relevant, the imperative to shift to a more sustainable and less carbon-intensive economic model becomes more urgent. The significant drivers vary by index, highlighting the different structural features of the various companies. The remainder of this paper is organized as follows. Section 2 provides a selection of the recent literature. Section 3 outlines the adopted methodology. Section 4 describes the dataset and their features. Section 5 summarizes the main results and concludes the paper. In the Appendix the reader can find additional tables and figures.

2 Literature Review

Capital markets are required to price the costs linked to running a sustainable business. Some of the questions to be answered are: what is the impact of a persistently unhappy workforce? How safe is a company's data; how vulnerable is the company to climate change? These questions are not captured in traditional financial analysis but are found to have a material financial impact. One way to deal with this is to turn towards ESG criteria to assess performance across non-financial factors. ESG criteria used to identify responsible investments are considered the three central factors to measure the sustainability and ethical impact of a company. ESG criteria are supposed to be represented by the assigned ESG score, which is supposed to help investors to deal with the unmeasured environmental, social, and governance topics. ESG scores are therefore used in literature to summarize the non-financial performance of a company providing disclosure regarding the sustainability of a company in its broadest sense. The relationship between a company's ESG score and its financial performance has been widely investigated in the last decade. Two main approaches deserve to be mentioned, the first studying the relationship between the ESG score and the financial performance of a company measured by its asset's rate of return; the second analyzing the link between ESG rating and structural variables as measured by balance sheet's data. The findings vary depending on a variety of factors, such as the industry, the firm's core activity, and the time horizon (Friede et al. (2015), Ye et al. (2021)).

Some studies identify this relationship as positive, while others conclude to a no significant, or even negative, relationship (Hartzmark and Sussman (2019), Jo and Na (2012), Kotsantonis and Serafeim (2019), Lourenço et al. (2012), Oikonomou et al. (2012), Xie et al. (2019)). For instance, Eccles et al. (2013) claim companies with higher ESG score show higher financial returns and are hardly affected by period of financial distress. Similarly, Flammer (2015), using a sample of S&P500 companies, find that companies characterized by higher ESG score exhibit higher market valuations. According to Clark et al. (2015), investors view a company's ESG as an indicator of its financial performance, since higher ESG score is associated with lower costs of capital.

Other studies exhibit a mixed or negative relationship between a company's ESG score and its financial performance. For example, Friede et al. (2015) conclude that according to their literature review the relationship is unclear, because it depends on the industry and region where the authors conduct the analysis. Likewise, Jo and Na (2012) find companies with higher environmental performance had lower stock returns.

D'Amato et al. (2023) develop a regression model to predict the firms' profitability using both balance sheet information and the ESG score. They use a ML approach to show that the ESG score has a significant effect in the operating profit of a company.

To the best of our knowledge, there has been no attempt to investigate how the ESG score may be explained by the company's structural features for the European listed stocks.

However, we believe by pinpointing areas for improvement, managers can enhance their firm's ESG performance and exploit the financial benefits.

3 Methodology

We study the relationship between the ESG score, which is the response or target variable, and a set of predictors represented by some financial statements ratios that will be described in Section 4. As highlighted by D'Amato et al. (2021), ratios are further informative than absolute values and provide a better characterization of companies, which helps to explain the ESG scores. We use machine learning to predict the Refinitiv ESG scores and identify the most important input variables to determine the prediction.

The supremacy of machine learning over more traditional methods, such as linear regression, mainly falls in revealing complex and hidden relationships. Among the machine learning algorithms proposed in the literature, the ensemble methods are acknowledged to provide better accuracy than individual techniques. Gradient Boosting Machines and Random Forests are ensemble algorithms widely used in several fields, especially when dealing with small or medium tabular data, where they revealed very powerful. In general, the decision of which algorithm to use depends on the specific task to perform and the characteristics of the data. Since our dataset has a large number of features compared to the number of observations and incorporates enough noise, we use the Random Forest algorithm, which is well-known to be less prone to overfitting, and more robust to noise and outliers than the Gradient Boosting Machine.

When dealing with machine learning tasks, the purpose is not only to find the most accurate estimation of the response variable but to identify the most meaningful variables to make predictions, leading to a deeper understanding of the problem under investigation. Therefore, we assess the importance of the input variable, and interpret the results through two well-known model-agnostic interpretation methods, Partial Dependence Plot, and feature interaction.

Random forest (RF). We summarize the RF algorithm's steps as follows. Firstly, it defines multiple samples from the training data using a bootstrap methodology to fit a weak learner on each one of them. Secondly, it computes the average of the weak learners' results to obtain an aggregate output. The way the RF selects its predictors represents the actual algorithm's strength and peculiarity. It chooses a random subset of predictors as candidates for the subdivision from the final set of predictors at each split. This way, the algorithm prevents the strongest predictors to become predominant in the splitting process. Also, the RF stochastically disrupts the learning system to discriminate trees and then combine their predictions through an aggregation technique Breiman (2001). Given a set of features, $X_1, X_2, ..., X_p$, in a certain predictor space, the RF estimator is defined as:

$$\hat{f}^{RF}(\mathbf{X}) = \frac{1}{B}\sum_{b=1}^{B}\hat{f}^{DT}(\mathbf{X}|b), \tag{3.1}$$

where B represents the number of bootstrap samples and $\hat{f}^{DT}(\mathbf{X}|b)$ the decision tree estimator over the $b \in B$ sample. The decision tree estimator $\hat{f}^{DT}(\mathbf{X})$ is obtained by recursively splitting the predictor space into J distinct and non-overlapping regions, $R_1, R_2, ..., R_J$, and is calculated as $\hat{f}^{DT}(\mathbf{X}) = \sum_{j \in J} \hat{y}_{R_j} \mathbb{1}_{\{\mathbf{x} \in R_j\}}$, where $\hat{y}_{R_j}$ is the response variable estimated by the mean of the observations falling in the same region R_j, and $\mathbb{1}_{\{.\}}$ is the indicator function.

Variable importance The general approach to show how the various input variables affect the response variable is to depict the variable importance plot based on a given measure. The literature suggests a few definitions of variable importance for regression trees (see, e.g., Louppe et al. (2013) and Biau and Scornet (2016)). We use the weighted impurity measure introduced by Breiman (2001, 2003), which estimates the relative impact of each input variable in predicting the response variable by summing up the weighted impurity decreases for all nodes of the tree where a given input variable is used, averaged over all the trees in the forest. Impurity is also known as the mean decrease in the Gini index.

Partial Dependence Plots (PDP). Introduced by Friedman (2001), this model agnostic tool shows the marginal effect of one or more features on the prediction of the target variable produced by the model. Let X_S be the feature of interest, and X_C the other variables considered in the model, with C the complement set of S. We compute the partial dependence (PD) function of f on X_S ($f_S^{PD}(x_S)$) as the marginal expectation over the features in set C:

$$f_S^{PD}(x_S) = \mathbb{E}_{X_C}\left[f(x_S, X_C)\right]. \tag{3.2}$$

We calculate the integral over the predictions weighted by the probability distribution of X_C ($P(X_C)$) to write the PD function of f on X_S, as follows:

$$f_S^{PD}(x_S) = \int f(x_S, X_C)dP(X_C). \tag{3.3}$$

However, one can estimate $f_S^{PD}(x_S)$ as $\hat{f}_S^{PD}(x_S) = \frac{1}{n}\sum_{i=1}^{n} f(x_S, x_{iC})$, where i represents the *i-th* observation.

Feature interaction. Once we read the principal marginal effects for the most important features through the PDP, we analyze the strength of the overall interaction of each single feature with the others. Friedman (2001) recurs to the H-statistic to define the interaction between a feature j and any other feature, which is defined as follows:

$$H_s^2 = \frac{\sum_{i=1}^{n}\left[\hat{f}(x^{(i)}) - PD_j(x_j^{(i)}) - PD_{-j}(x_{-j}^{(i)})\right]^2}{\sum_{i=1}^{n}\hat{f}^2(x^{(i)})}. \tag{3.4}$$

H_s^2 ranges in the interval $[0, 1]$, where the null value indicates the absence of interaction, while 1 full interaction. It shows the portion of the estimated target variable's variance explained by each interaction.

4 Empirical Analysis

4.1 Data description

We consider the 416 listed companies constituent the eight major European stock indexes over the sample period 2012-2021. Data regarding the ESG scores and the single E, S, G pillars, together with the balance sheet data are retrieved using the provider Refinitiv (Refinitiv, 2022b). The eight Stock Indexes are:

- BEL 20 (Euronext Brussels, Belgium)
- CAC 40 (Euronext Paris, France)
- DAX40 (Frankfurt Stock Exchange, Germany)
- MIB 40 (MIB) (Borsa Italiana, Italy)
- AEX 25 (Euronext Amsterdam, the Netherlands)
- IBEX 35 (Bolsa de Madrid, Spain)
- OMX Stockholm 30 (Stockholm Stock Exchange, Sweden)
- FTSE 100 (London Stock Exchange, United Kingdom)

For each company included in the Index, we retrieve the following accounting, market and ESG data:

- Return On Equity (*ROE*)

- Return On Investment (ROI)
- Return On Asset (ROA)
- Solvency Ratio (SR)
- Price to Earnings (P/E)
- Dividend per Share (DPS)
- Carbon Intensity (CI)= (Total Net Carbon Dioxide emissions, tons)/(Revenues, million USD)
- ESG score (ESG) (min. 0; max. 100)
- E score (E)
- S score (S)
- G score (G)
- ESG Controversies Score ($ESGCS$), that indicates the company's exposure to ESG controversies and negative events reflected in global media (Refinitiv, 2022b).

We also estimate some specific ratios as:

- EBIT margin ($EBITtoRev$) = (Earnings Before Interest and Taxes (EBIT))/ Revenues
- Asset Turnover (AT) = (Revenues)/(Total Assets)
- Net Profit Margin (NPM) = Net Income/Revenues
- Liquidity Ratio (LR) = Stock of highly liquid asset/Total net cash outflows
- ($MCtoEBIT$) = Market Capitalization/EBIT
- Size of a company ($Size$) = $\log 10(TotalAssets)$.

The data cleaning process[2] leads to a final dataset of 410 firms classified as reported in Fig. 1 and Fig. 2. The large part of the sample is composed by UK firms (138; 34%), followed by German (52; 12%), French (49; 11%), Italy (39; 9%) which represent the 57% of the entire sample. Each Index shows a different distribution of companies among the various economic sectors. The French index (CAC 40) comprises mostly companies involved in the Consumer Cyclical business

[2]We remove the companies for which we lack some of the data.

(31%), Industrials (20%) and Technologies (14%), unlike Germany (DAX-40), where Basic Materials (19%), Consumer Cyclicals (19%) and Healthcare (14%) play the main roles. Consumer Cyclicals companies rely heavily on the business cycle and economic conditions (i.e. industries such as automotive, housing, entertainment, and retail), and companies involved in the automotive sectors have been modifying the business model from the beginning of 2000. This may explain the presence of companies with higher ESG scores in this sector. The Basic Materials sector includes companies involved in the discovery, development, and processing of raw materials (oil, stone, and gold). The companies operating in this sector are required to be more active toward sustainable goals and show ESG scores more dynamics than others. Utilities play a major role in Spain (IBEX) (18%) and Italy (MIB) (15%), in these countries two large companies as Enel and Iberdrola have been involved in the sustainable goals earlier than other companies and report ESG scores higher than other from 2013. The Financial sector is predominant within the UK and Italy, companies in this sector have been recently challenged by the regulator (ECB) to comply with the ESG risk measurement and management. Technology is the main sector in the Netherlands (23%), in Belgium (19%) and Sweden (22%).

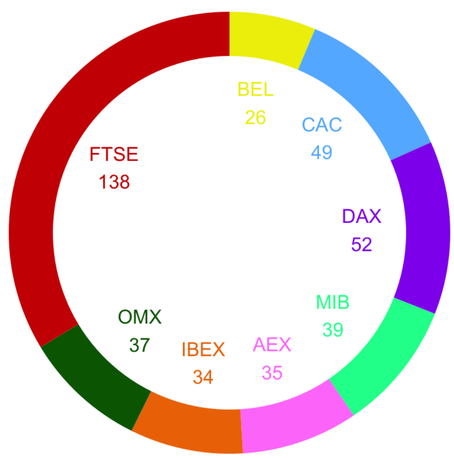


Figure 1: Number of firms in the sample by stock index.

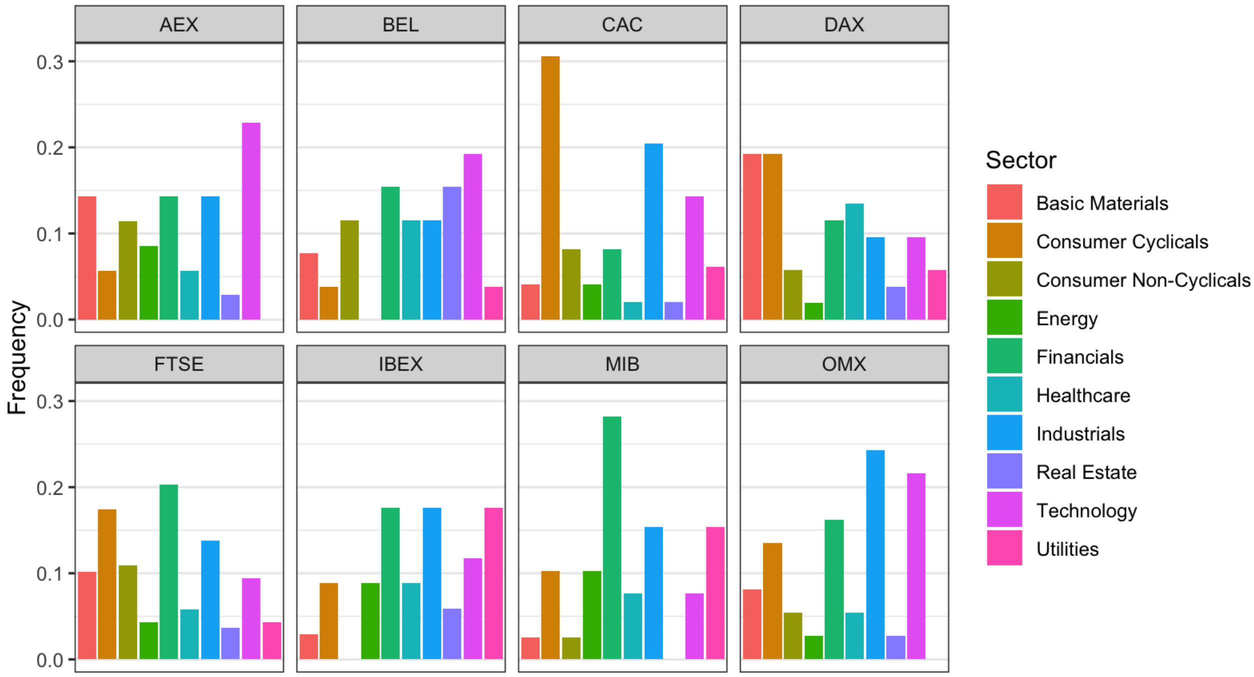


Figure 2: Economic sectors by stock index (TRBC Economic Sector classification, Refinitiv).

UK and Italy are by far the markets with largest trading volumes as it is shown in Figure 3, until 2016 the trading volumes in UK and Italy exceeded 150 billion per year while the other markets were all around 50 billions. After 2017 The UK market is the only one showing trading volumes larger than USD 200 billion. If we look at the distribution of volumes between the various indices, UK, Italy and France cover 80% of the entire European market as it is reported in Table 1. It is interesting to report a change occurred during the sample period between the volume share in UK (31% in 2013 increasing to 46% in 2022) balanced by the reduction observed in the volume share of the Italian index which moved from 40% in 2013 to 24% in 2022. The France market did not report particular changes over the period, the other markets show a steady share of the trading volumes over the sample period, Spain (10%) and Germany (4%) Sweden and Belgium represent steadily the 1% of the trading volume.

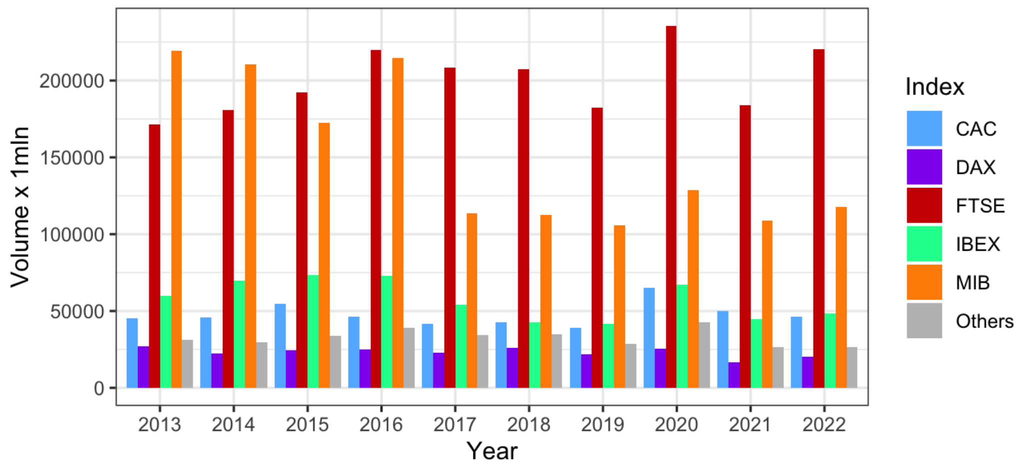


Figure 3: European exchanges yearly volumes in USD mln. Others contains the sum of AEX, BEL, and OMX volumes.

Year	AEX	FTSE	CAC	OMX	MIB	IBEX	BEL	DAX
2013	0.05	0.31	0.08	0.00	0.40	0.11	0.01	0.05
2014	0.04	0.32	0.08	0.01	0.38	0.12	0.01	0.04
2015	0.05	0.35	0.10	0.01	0.31	0.13	0.01	0.04
2016	0.04	0.36	0.08	0.01	0.35	0.12	0.01	0.04
2017	0.05	0.44	0.09	0.01	0.24	0.11	0.02	0.05
2018	0.05	0.45	0.09	0.01	0.24	0.09	0.02	0.06
2019	0.05	0.44	0.09	0.01	0.25	0.10	0.01	0.05
2020	0.05	0.42	0.12	0.01	0.23	0.12	0.02	0.05
2021	0.05	0.43	0.12	0.01	0.25	0.10	0.01	0.04
2022	0.04	0.46	0.10	0.01	0.24	0.10	0.01	0.04

Table 1: Volumes proportion per index per year computed as index yearly total volume over the annual total volume.

Figure 4 reports the pairwise correlations between the variables during the sample period. We do not find strong associations within variables, except among those self-evident as the ESG score and its components, E, S, G, where we found correlations higher than 0.5; and the corporate ratios as the ROA, ROE, and the ROI, which show correlations higher than 0.8. Also, no significant differences of correlations are found within each index or by year. The *ROA*, *ROE*, and *ROI* results strongly and positively related (66-85%). The *NPM* is linearly associated only with the LR (58%). The highest correlation is found between *Size*, *AT*, *ESGC* and the *ESG* scores. Interestingly, the *CI* is uncorrelated with the *E* score, underlying the lack of coherence among the currently available environmental metrics (see Popescu et al. (2021), Berg et al. (2022), among others).

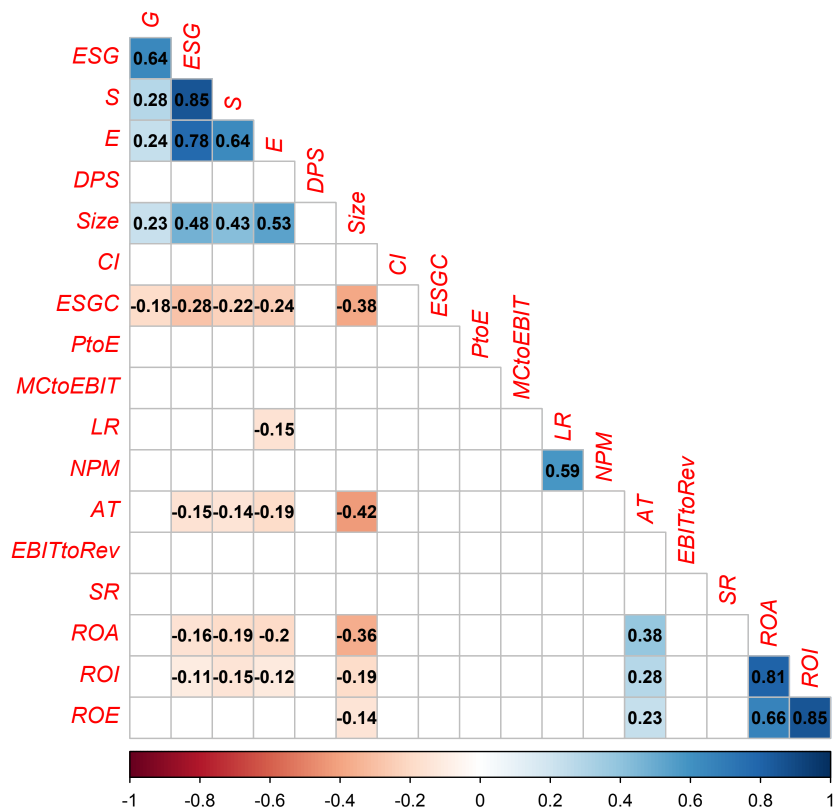


Figure 4: Variables correlation (2012-2021).

The average ESG scores for each Index by year is reported in Figure 5. Figure 6 reports the same representation for each ESG component, E, S and G. The average value of the ESG score shows an almost constant rise over the years for all the indices. The French (CAC) and the German (DAX) index peak over the others, reaching the highest score in 2021 (CAC=79.5; DAX=79.81), they are the most Sustainable Development Goals (SDGs) aligned countries among those considered in the analysis (SDSN Europe, 2022; World Economic Forum, 2020). The other indices exhibit a steady growth over the years with the exception of the Netherlands (AEX), which shows a faster pace after 2017 and reaches the second higher score in 2021, and the Belgian (BEL) index which starts in 2013 from a very low score (45.99) and grows faster over the years. The ESG score in Belgium shows also a larger fluctuation with a remarkable increase after 2017 which reduces the average ESG score gap and could be a consequence of the reception of the EU Non-Financial Reporting Directive (European Union, 2014) in this country where most of the listed companies operate in the technology sector, followed by real estate and financials. The early adoption of this directive could be the reason why France and Germany show superior *ESG* performance (Jones Day, 2021). Looking at the dynamics of the single components it is worth noticing that the only score which shows a steady growth over the years is the S score, with France and Germany still leading the pace and UK and Belgium showing the lowest scores. The E scores, show an almost constant level from 2013 until 2017 for each country, and starts increasing only in the last few years, with Belgian companies becoming quite active in the Environment component after 2017. The effective global action to address the impacts and future risks of the climate changes, introduced with the Kyoto protocol, entered into force in 2005 and after the Doha amendment in 2012, had become a major challenge for European governments' aiming to ensure a net-zero transition by 2050. This brought a growing number of corporates, financial institutions and institutional investors to make efforts to set climate transition plans to achieve net-zero emissions already at the beginning of the second decade of the new millennium. Most European companies started to dramatically reduce emissions, shift away from carbon-intensive activities and promote green growth, generating a level of the E score quite high already in 2013 that remained almost unchanged until 2016. The Paris agreement of 2015 had forced governments and companies to put more effort in sustainability issues and on climate change risks. The Governance component, unlike the E score, reports lower scores for all the indices during the period 2013-2017 and a quite slow, or none, growth, only after 2017 it increases for all the countries. Germany results the most active in the Governance pillar, the G score moves from 60 in 2013 to 85 in 2021, while France, very active for the E and the S component, shows a different pattern for the G component, as Spain which reports the lowest G score, just slightly above the Belgian index. The Belgian index shows the lowest E, S, and G values each year. The dynamics of the various E, S, G components depends on the different economic sector distribution. For instance, in countries such as France, where most companies operate in the Consumer cyclicals sector (automotive, retail), attention toward climate change started in the first decade of the new century. This explains the high E score already in 2013, while attention toward Governance was imposed only after the Paris Agreement in 2015. We find the only exception in the period 2019-2020, where the UK and BEL indexes assume similar values, both largely different from the others.

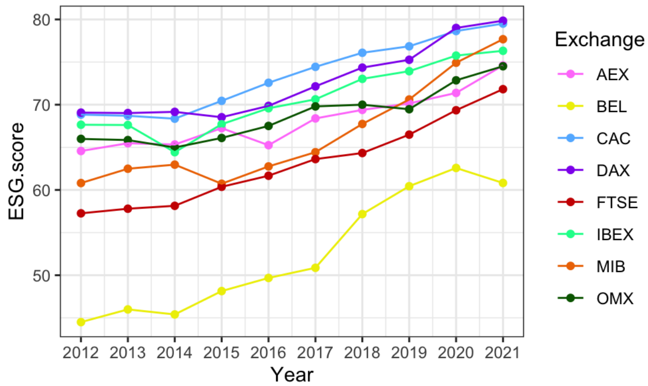


Figure 5: Average ESG score per year by stock index.

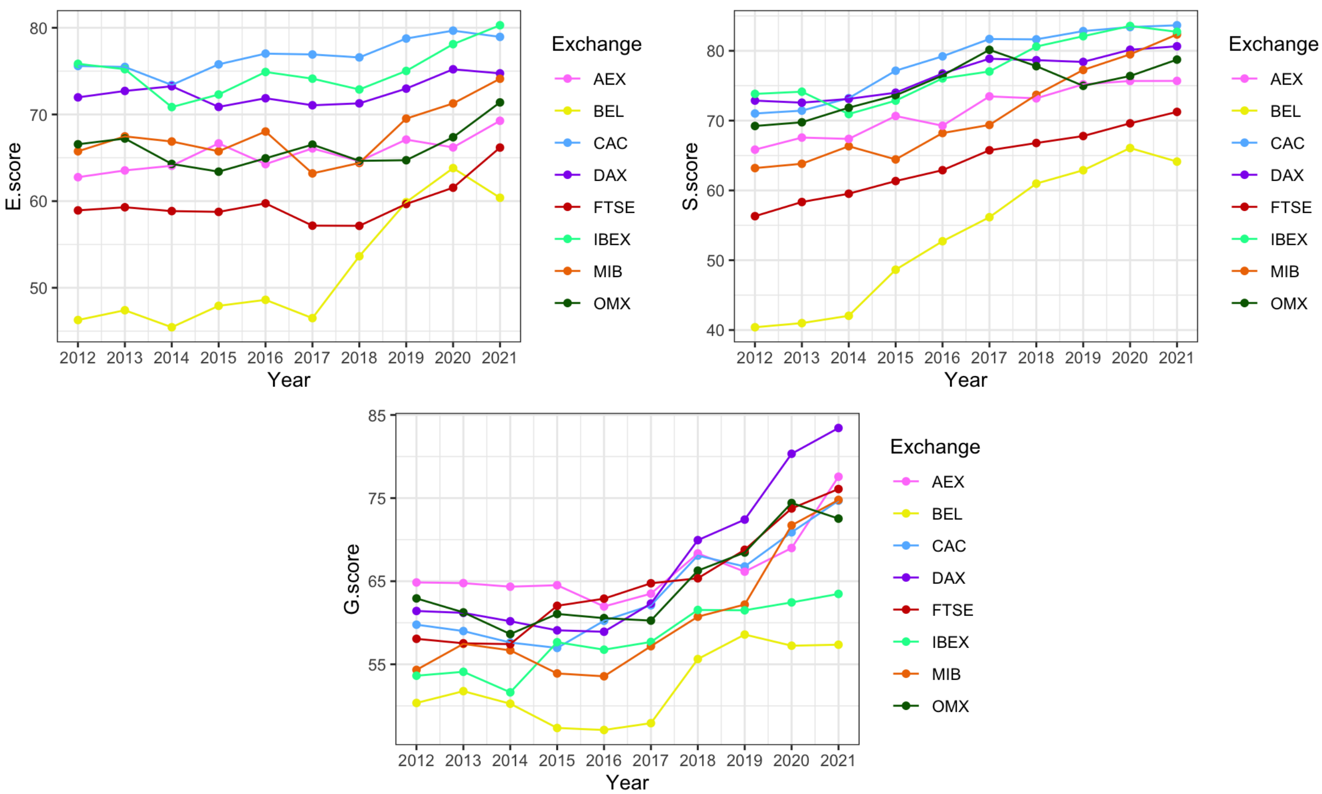


Figure 6: Average E (top left panel), S (top right panel), and G (bottom panel) score per year by stock index.

The Controversy score was introduced as a tool to capture harmful businesses practices, it aims to highlight monopolistic behaviors, fraud and harassment in the workplace, and environmental incidents. The ESG controversies score is used to get a rounded and comprehensive scoring of a company's ESG performance and contributes to the creation of the ESG Combined score. The main objective of the ESG combined score is to discount the ESG performance score based on negative media stories. We use the Refinitiv controversies score, based on 23 ESG controversy topics, with recent controversies reflected in the latest complete period (for further details, refer to (?)). When companies are involved in no ESG controversies, the ESGC score is equal to 100, otherwise it is estimated as the average of the severity of each controversy and the number of controversies occurring in the fiscal year. The Figure 7 reports the variation in the average controversy score *ESGC* and the combined score *ESGComb* for each index. It is interesting to notice that the French and German indexes show the lowest average *ESGC* score over the years, evidencing the occurrence of several controversies. This offsets the ESG performance of these two indices and result in a lower level of the *ESGComb*. A couple of events may explain the controversies scores in Germany and France. The France government recently admitted to being in late in alignment with the SDGs. In 2022 Environmental groups sued TotalEnergies over climate marketing claims, and in 2019 Investigation alleges Lactalis breached environmental regulations at several plants and released milk derivatives that killed fish. Differently, recent German history is characterized by several environmental and ESG scandals. Among the largest ones, we find the Dieselgate in 2015, when a large automotive company was found guilty of *greenwashing* (BBC, 2015; European Commission, 2017) or, more recently, the 2021 ESG scandal which saw the Deutsche Bank AG's DWS Group accused of exaggerated the sustainable features of the ESG-labeled financial products offered (Bloomberg (2021), Financial Times (2022)). This result underlines the importance of accompanying an ESG-based analysis with the relative *ESGC* score. Figure 7 reports the average ESG combined score, showing that in terms of Combined ESG score Spain and Sweden are the best-performing countries. Even Belgium has a higher score in 2021. The Carbon Intensity (*CI*) is a measure of how clean our electricity is. It refers to how many grams of carbon dioxide (CO2) are released to produce a kilowatt hour (kWh) of electricity. The non-financial information is becoming more important when assessing investment risk and opportunities so listed companies are more and more required to report the non-financial information. Large part of the companies in our sample report on CO_2 emissions data, for the companies missing these data Refinitiv developed sophisticated carbon data and provides an estimate for it. We use the Carbon intensity/Revenues ratio which measures the ton of CO_2 by USD. The average CI by country over the sample period is reported in Figure 7. (Total Net Carbon Dioxide emissions, tons)/(Revenues, million USD) The Belgian, French, Spanish, Swedish, and the UK indexes do not show high volatility of CIs over the years, while the German, the Italian, and the Dutch ones broadly fluctuating with the DAX and the MIB showing a steady downslope trend from 2017. The different economic sectors' composition of the sampled indices makes it difficult, and maybe misleading, to compare their average *CI*. The

OMX shows the lowest average CI values, this is mainly due to the leading role of hydropower plants to produce electricity in Scandinavian countries (SDSN Europe (2022)).

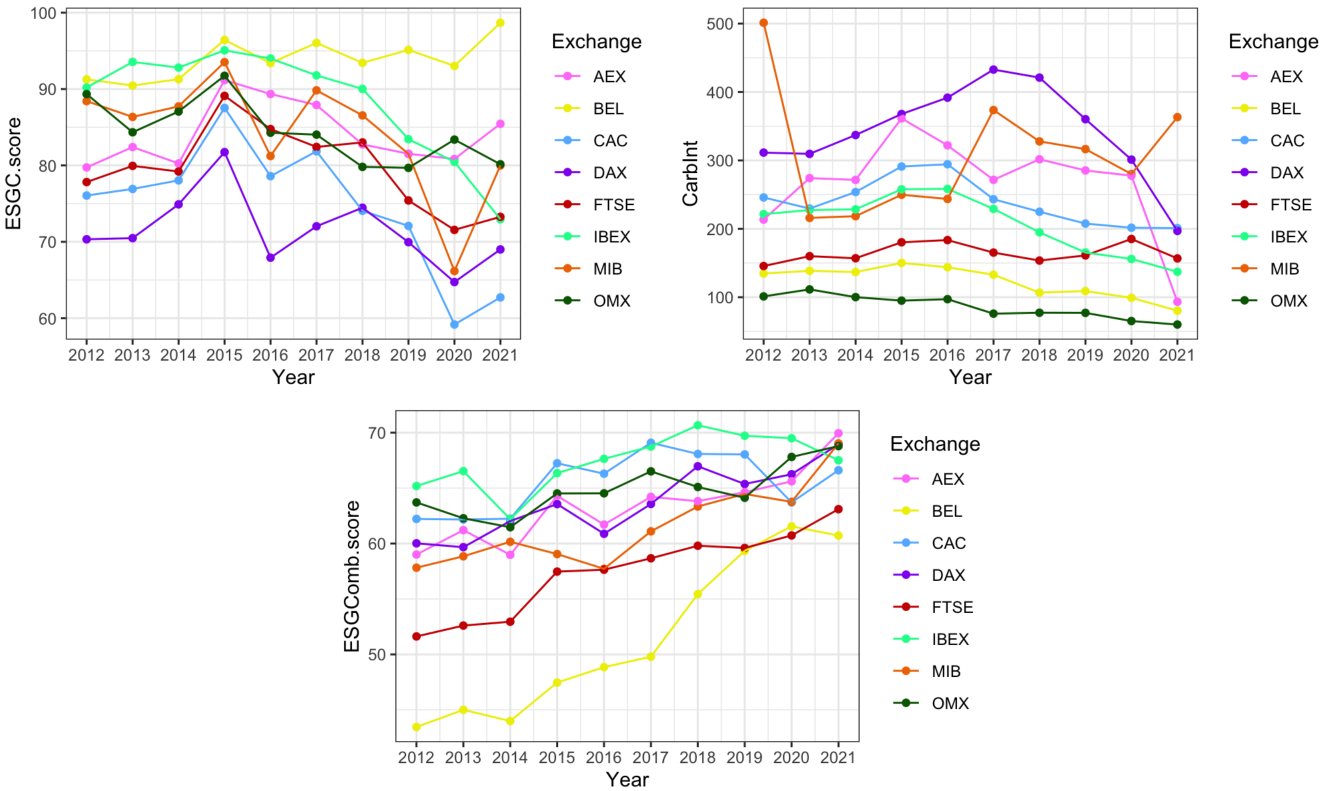


Figure 7: Average Controversies score (first row, left panel), Average Carbon Intensity (first row, right panel), and Average ESG Combined Score (second row) per year by stock index (2012-2021).

4.2 ESG score prediction

We apply the RF algorithm to the following regression model:

$$ESG \sim DPS + Size + CI + ESGC + PtoE + MCtoEBIT + LR + NPM + AT+ \\ + EBITtoREV + SR + ROA + ROI + ROE \tag{4.1}$$

We evaluate the model's ability to predict the ESG score through two conventional error measures, the root mean square error (RMSE) and the mean absolute percentage error (MAPE). Their values for the train set (70% of the data) and test set (30% of the data) are shown in Table 2. Since MAPE provides a percentage error, it allows for an easy understanding of the goodness of the error value. Though the MAPE score depends on the dataset, a general rule to be followed is that a MAPE value lower than 10% is considered a very good value, and a MAPE value in the range of 10%-20%

is considered a good value. Therefore, the performances of our model are very good for AEX and DAX and good for the other indexes in both the train and test samples.

Index	RMSE train	MAPE train	RMSE test	MAPE test
AEX	7.45	8.67%	6.31	6.78%
BEL	8.76	11.95%	8.91	12.57%
CAC	8.79	10.63%	9.14	11.16%
DAX	7.82	9.22%	7.59	8.27%
FTSE	9.17	13.55%	9.93	14.61%
MIB	9.56	12.00%	9.68	12.58%
IBEX	9.47	12.06%	8.93	12.03%
OMX	9.38	10.58%	9.68	12.10%

Table 2: Model performance by stock index. Target variable: ESG score.

We show in Fig. 8 (indexes: AEX, DAX, FTSE, MIB) and Fig. 9 (indexes: CAC, IBEX, BEL, OMX) the variable importance found for each stock index. From Fig. 8, we observe that for Germany, The Netherlands, Italy and the UK, the most important variable in explaining the ESG score is the firm's size measured as a function of the total assets (it ranks first in 6 out of 8 stock indexes). In Germany, many companies operate in the consumer cyclicals and basic materials sector, the total assets of these companies provide resources to invest in a more sustainable business model, translating into higher ESG scores. In Italy and UK, most of the companies operate in the Financial sector, which has been especially forced by the regulator to be active in responsible business. The listed financial companies are mainly large. The second important variable for these four indices is Carbon Intensity, so companies having a very volatile number of emissions by revenue are also the ones who actively operate to reduce emissions. It is interesting to notice that, for the UK, *Size* and Carbon Intensity represent by far the most important variables; the rest of the variables, Controversial score, Solvency ratios, etc., provide a small contribution, unlike Germany or Italy where, together with the Carbon Intensity, the *NPM*, *AT*, and *SR* contribute to explain the ESG volatility. From Fig. 9, we find that also for France and Spain, the most influential variable in describing the ESG score is the firm's size, while we find *AT* for Belgium and *EBITtoREV* for Sweden.

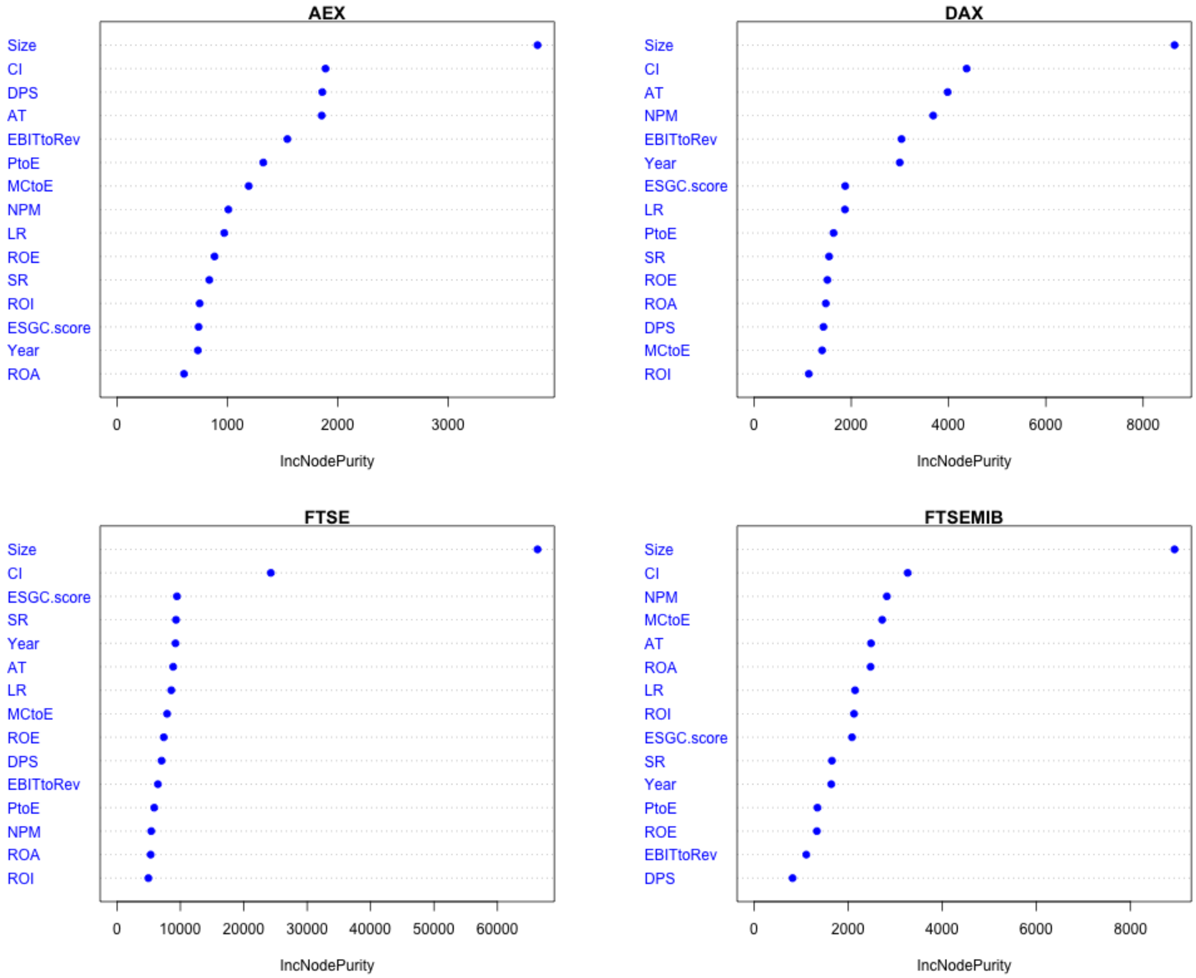


Figure 8: Variable Importance by stock index (indexes: AEX, DAX, FTSE, MIB). Target variable: ESG score.

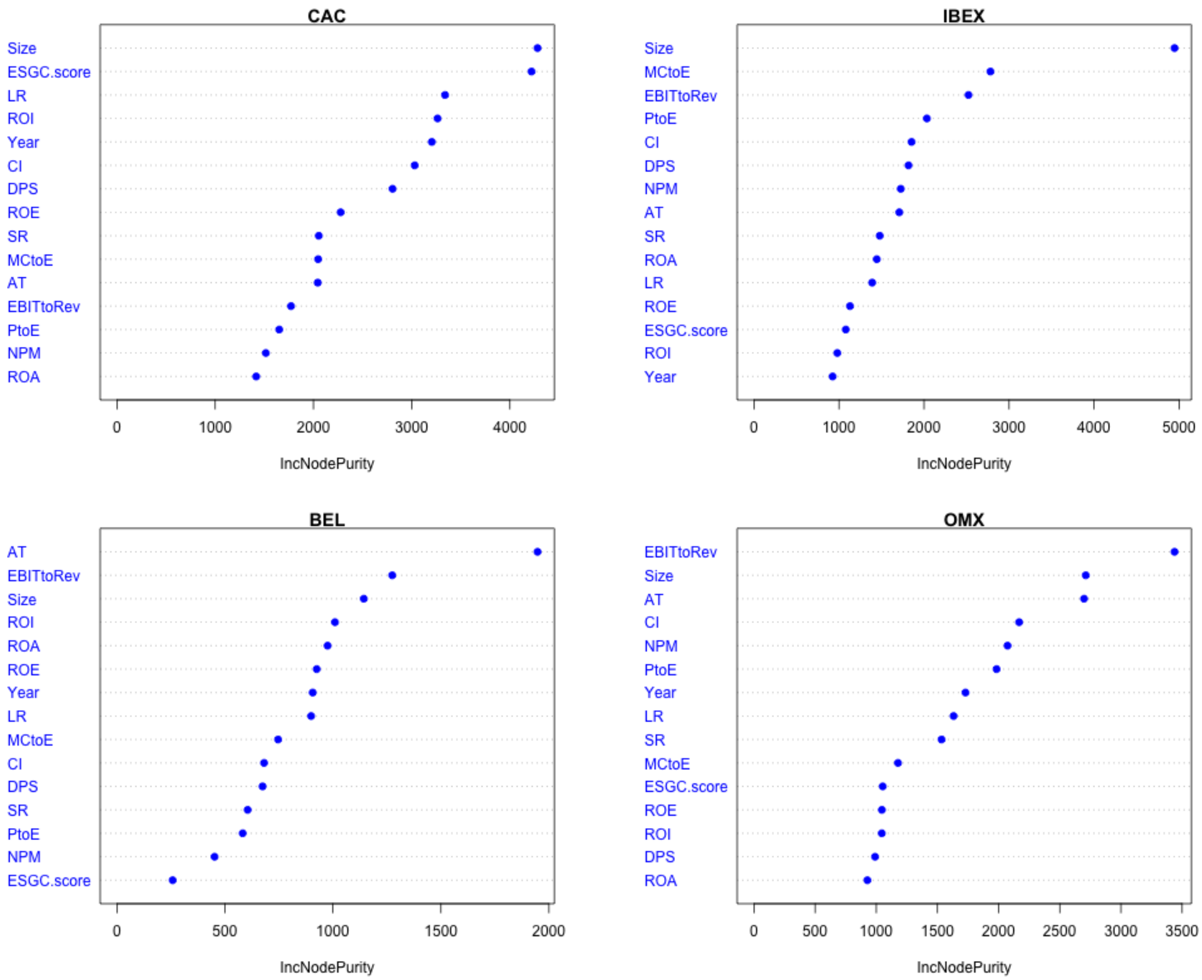


Figure 9: Variable Importance by stock index (indexes: CAC, IBEX, BEL, OMX). Target variable: ESG score.

To better understand the behavior of each influential predictor within the various stock indexes and find common patterns, we analyze the PDP. In Fig. 10, we depict the PDP of the *Size* predictor in the stock indexes where it ranks first (AEX, CAC, DAX, FTSE, MIB and IBEX). The PDP of *Size* highlights that the six stock indexes have a common behavior, showing an increasing nonlinear trend with jumps. A higher jump is observed for *Size* values of about 10.1-10.2 (13-16 billion euros). Therefore, companies with a size greater than 10.1-10.2 have a noticeably higher ESG score than smaller companies. For example, in the FTSE, companies with size 9 have an ESG score of 60, while companies with size 10.2 have an ESG score of 70. We can speculate that there is a size-threshold effect on ESG rating. Indeed, larger companies commonly have additional resources to develop more refined sustainability policies that usually lead to achieving higher ESG scores than their peers.

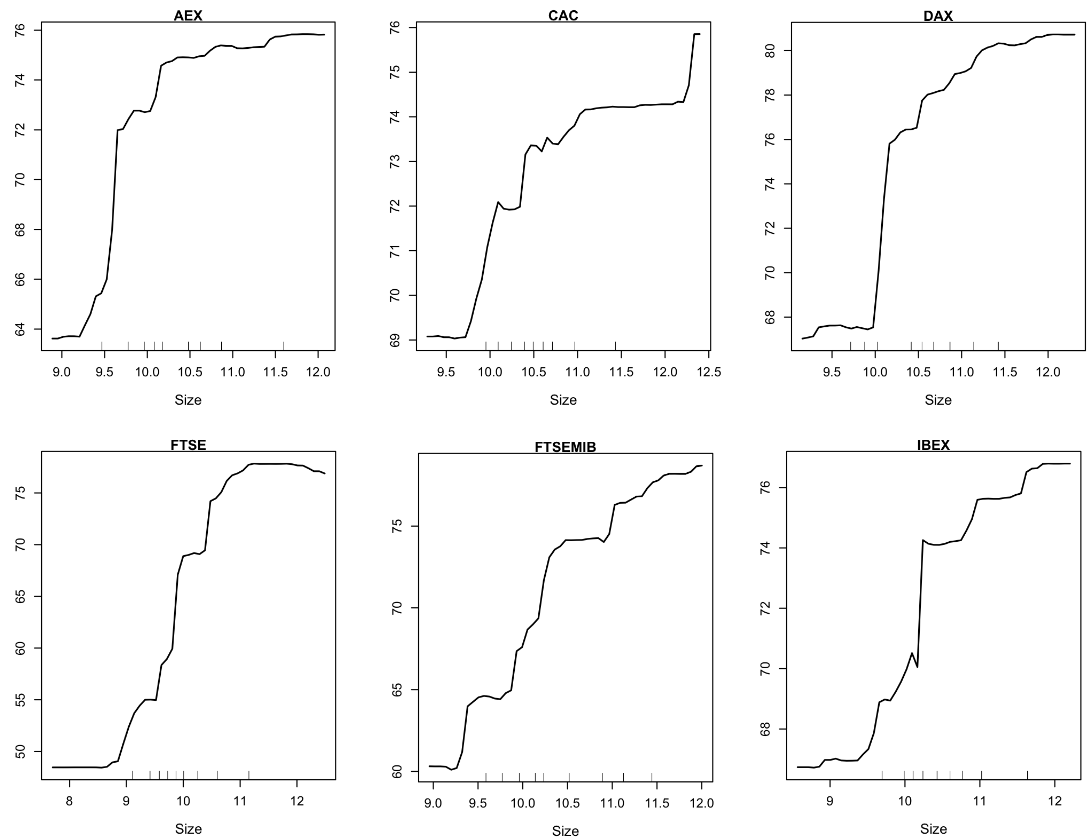


Figure 10: PDP. Target variable: ESG score. Variable: *Size*.

Fig. 11 shows the PDP of the Carbon Intensity predictor in those stock indexes where it ranks among the top three predictors. We observe that *CI* (we remind that it is the ratio between the total net carbon dioxide emissions in tonnes and revenues expressed in million USD) shows a growing trend and reaches a plateau around the value of 500 in the AEX and MIB and 1000 in the FTSE index. Companies having low *CI* show a much lower ESG score than those with high *CI*. To better understand the relationship between the carbon intensity and ESG rating in these four stock exchanges we display their values referred to the entire dataset in a scatterplot, Fig. 18, reported in the Appendix. Since our data is characterized by a large value range of carbon intensity, we use a logarithmic scale, which is sound for displaying data more compactly. We focus on indexes where *CI* is one of the most important predictors, the AEX, DAX, FTSE, and MIB. For the AEX, MIB, and DAX, we observe a non-linear relationship between the ESG score and Carbon Intensity. A peculiar aspect emerges, especially for AEX and DAX, within the range of ESG scores between 60 and 80, where most observations are typically concentrated. This range may represent a 'critical zone' where the dynamics between *CI* and the ESG score could significantly influence investment decisions and sustainability assessments for investors and companies. Another fascinating behavior emerges from the scatterplot of the DAX index, where for an ESG score higher than 80, we observe a negative relationship with *CI*. The lower the Carbon Intensity, the higher the ESG score. We find a similar, but less powerful, behavior for the MIB index in the same ESG score range. Instead, in the case of the AEX and FTSE indexes, the relationship between ESG score (higher than 75) and *CI* is weakly positive. However, it is noteworthy that scatterplots show an extensive cloud of points, which might suggest a certain degree of noise or variability in the relationship.

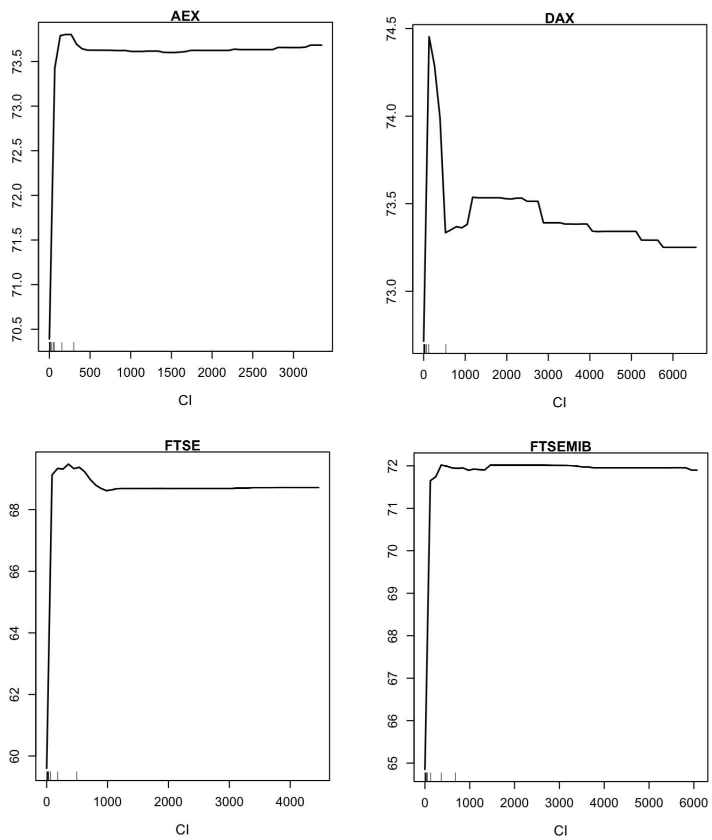


Figure 11: PDP. Target variable: ESG score. Variable: *CI*.

Finally, Fig. 12 and 13 depict the PDP of the AT and $EBITtoREV$ predictors, respectively, in those stock indexes where they ranks among the top three predictors. We remind that the asset turnover ratio AT assesses the efficiency with which a company uses its assets to generate revenues. A company with a high AT is more efficient than a company with a lower AT value. The ESG rating shows higher values when AT is between 0.75 and 1.25 in the DAX and OMX indices (Fig. 12). The PDPs of AT of these two stock indexes show similar behavior. For AT values higher than about 1.5, the ESG score falls from 75 to 72 in the DAX index and from 72.5 to 69 in the OMX index. Different is the case of the BEL index, where the highest ESG score values are observed for AT approximately equal to 0.7, then remarkably decrease for AT values higher than 1. Regarding the PDP of *EBITtoREV* (Fig. 13), we can observe different shapes for the three stock indexes (BEL, IBEX, and OMX), which have this feature among the top three as the level of importance in predicting the ESG score. In the case of the BEL index, the ESG score is quite stable for *EBITtoREV* values higher than 0.15, while for values lower than 0.15, it is first decreasing and then increasing. In the case of the IBEX and OMX indices, the ESG score grows by *EBITtoREV* values between 0 and 0.2 in the IBEX and 0.1 in the OMX and then decreases (more quickly in the case of the IBEX).

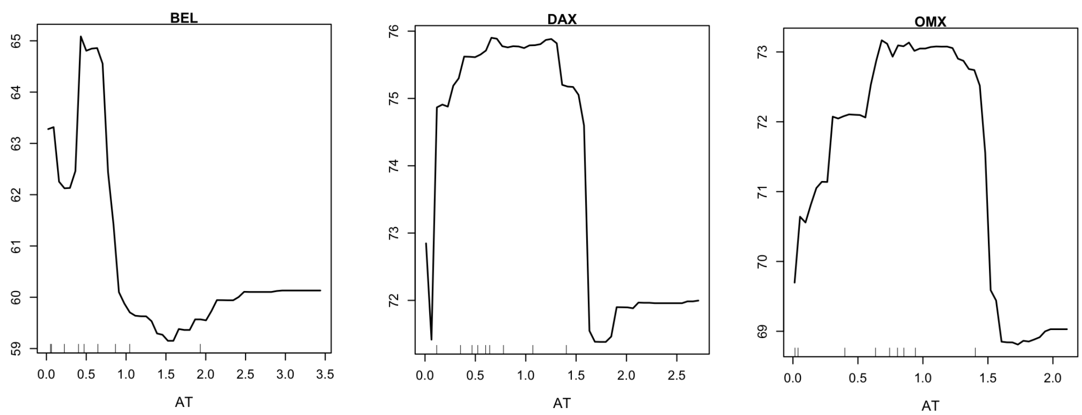


Figure 12: PDP. Target variable: ESG score. Variable: *AT*.

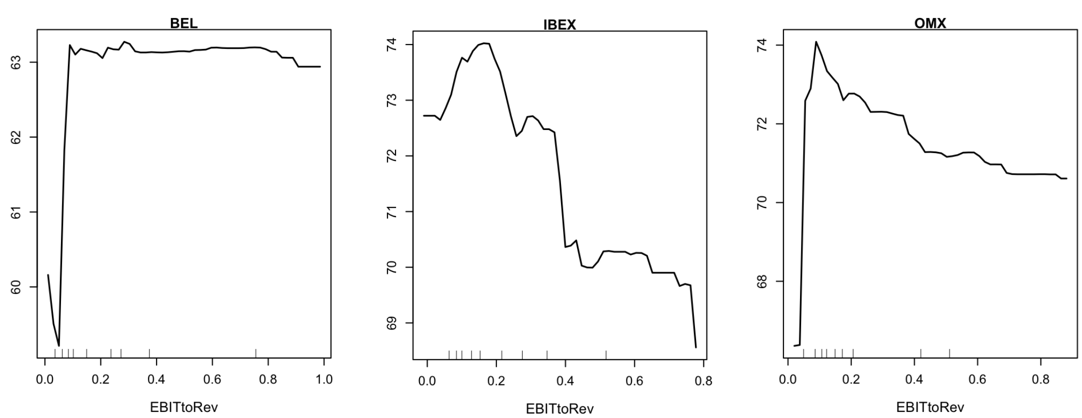
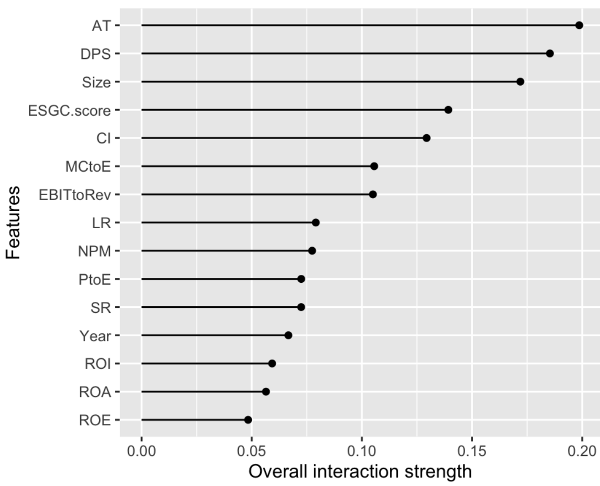
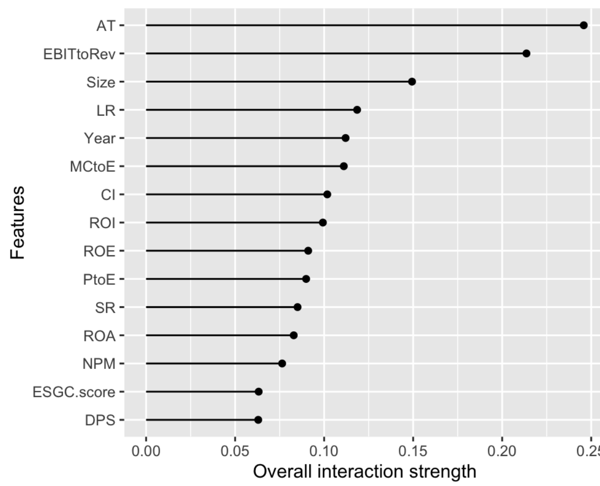


Figure 13: PDP. Target variable: ESG score. Variable: *EBITtoREV*.

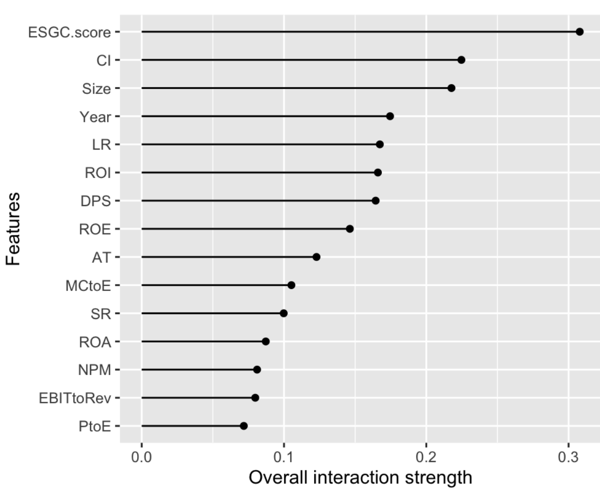
In Fig. 14 (indexes: AEX, BEL, CAC, DAX) and Fig. 15 (indexes: FTSE, MIB, IBEX, OMX), we show the overall feature interaction, i.e., how strong the interaction of a feature with all the others is in predicting the ESG score within each stock index. Feature interaction assesses the proportion of the variance of the estimated ESG score, which is explained by the interaction. We find that the interaction of the firm's *size* with the other features explains about 30% of the variance of the estimated ESG score in the DAX, FTSE, MIB, and IBEX indexes, and slightly more than 20% in the CAC index. The overall interactions of the *ESGC* score feature explain more than 30% of the variance in the CAC index, while the *AT* interactions account for 25% in the BEL index and about 20% in the AEX and OMX indexes. The interactions of *CI* are also meaningful as they explain more than 30% of the variance in the FTSE and more than 20% in the CAC and MIB indexes.



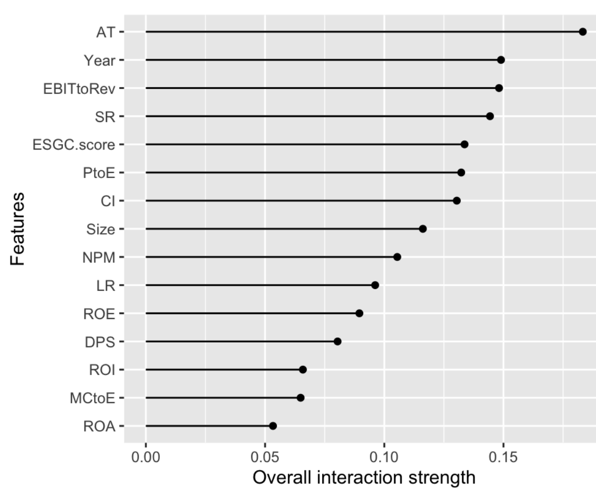
(a) AEX



(b) BEL

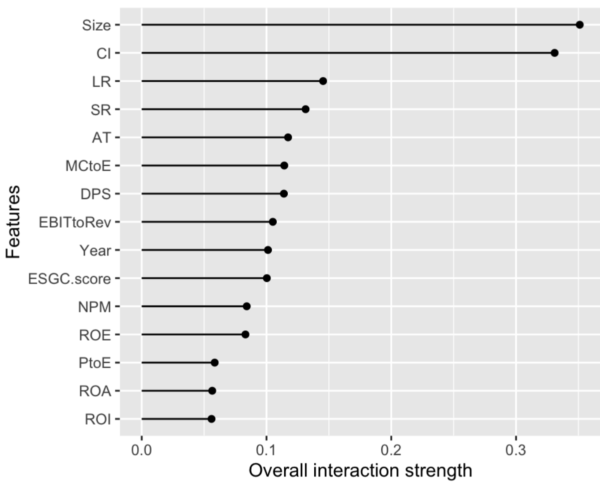


(c) CAC

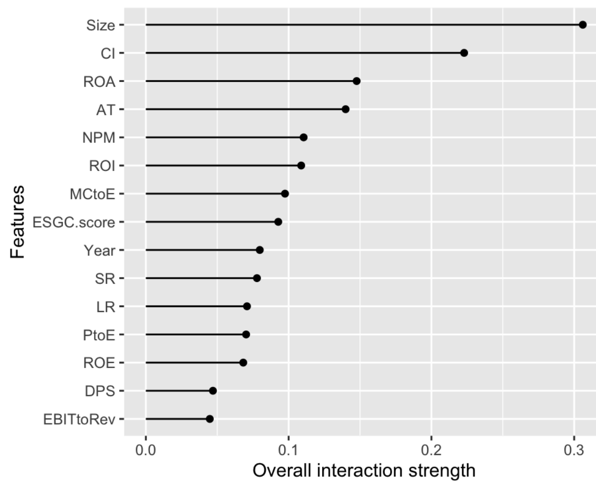


(d) OMX

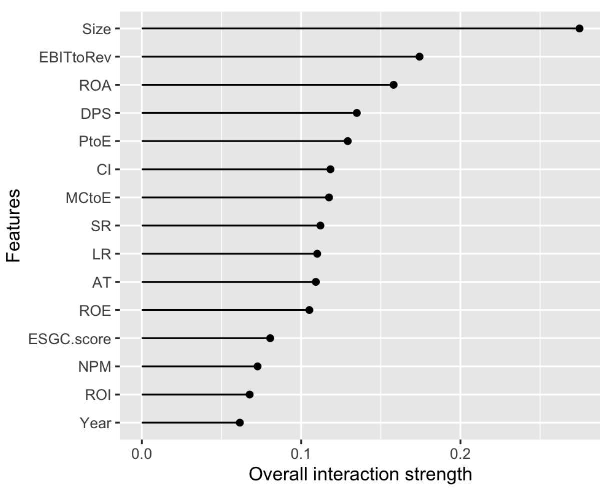
Figure 14: Interactions. Target variable: ESG score. Indexes: AEX (a), BEL (b), CAC (c), OMX (d).



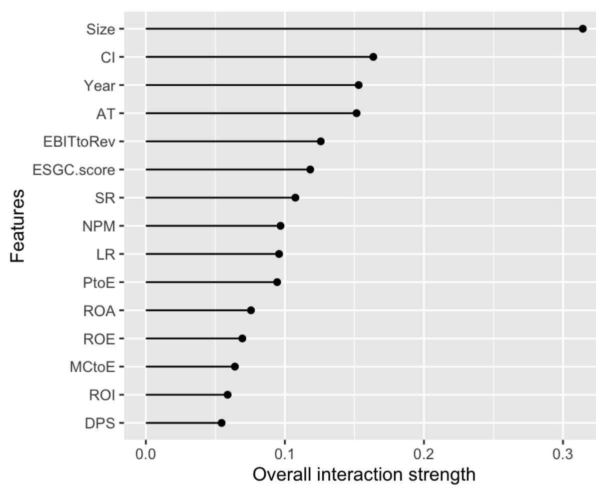
(a) FTSE



(b) MIB



(c) IBEX



(d) DAX

Figure 15: Interactions. Target variable: ESG score. Indexes: FTSE (a), MIB (b), IBEX (c), OMX (d).

To visualize the goodness of the prediction, we illustrate in Fig. 16 (indexes: AEX, BEL, CAC, DAX) and Fig. 17 (indexes: FTSE, MIB, IBEX, OMX) the ESG score prediction on the test sample compared to the observed values. The width of the red segment provides the extent of the estimation error produced by the model. This representation enables us to detect which point observations the model underestimates or overestimates. We generally observe that the model tends to overestimate ESG scores lower than about 70 and underestimate ESG scores higher than about 70. This behavior is less evident for the FTSE index.

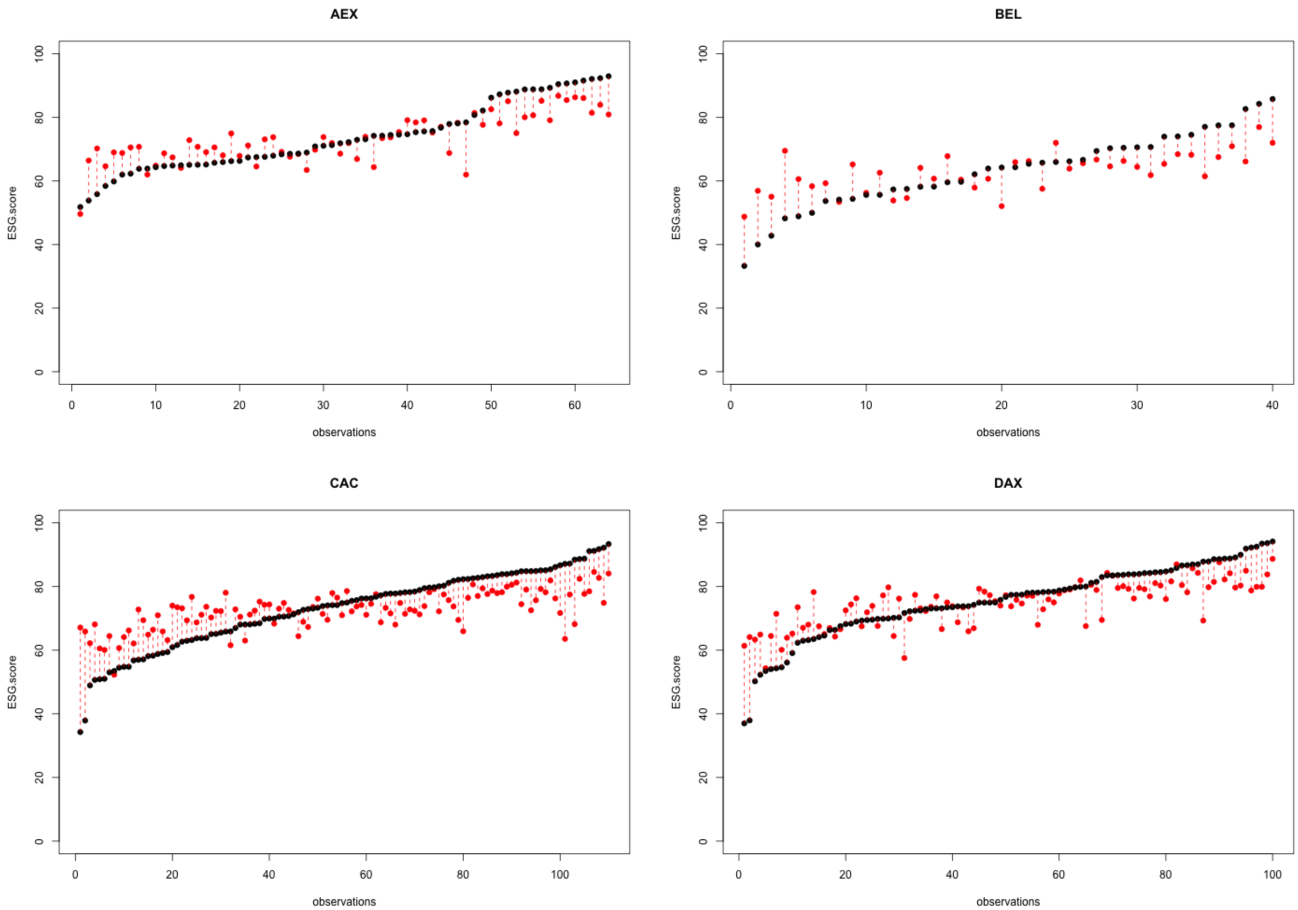


Figure 16: Predicted values (red dots) vs. observed values (black dots). Target variable: ESG score. Indexes: AEX, BEL, CAC, DAX.

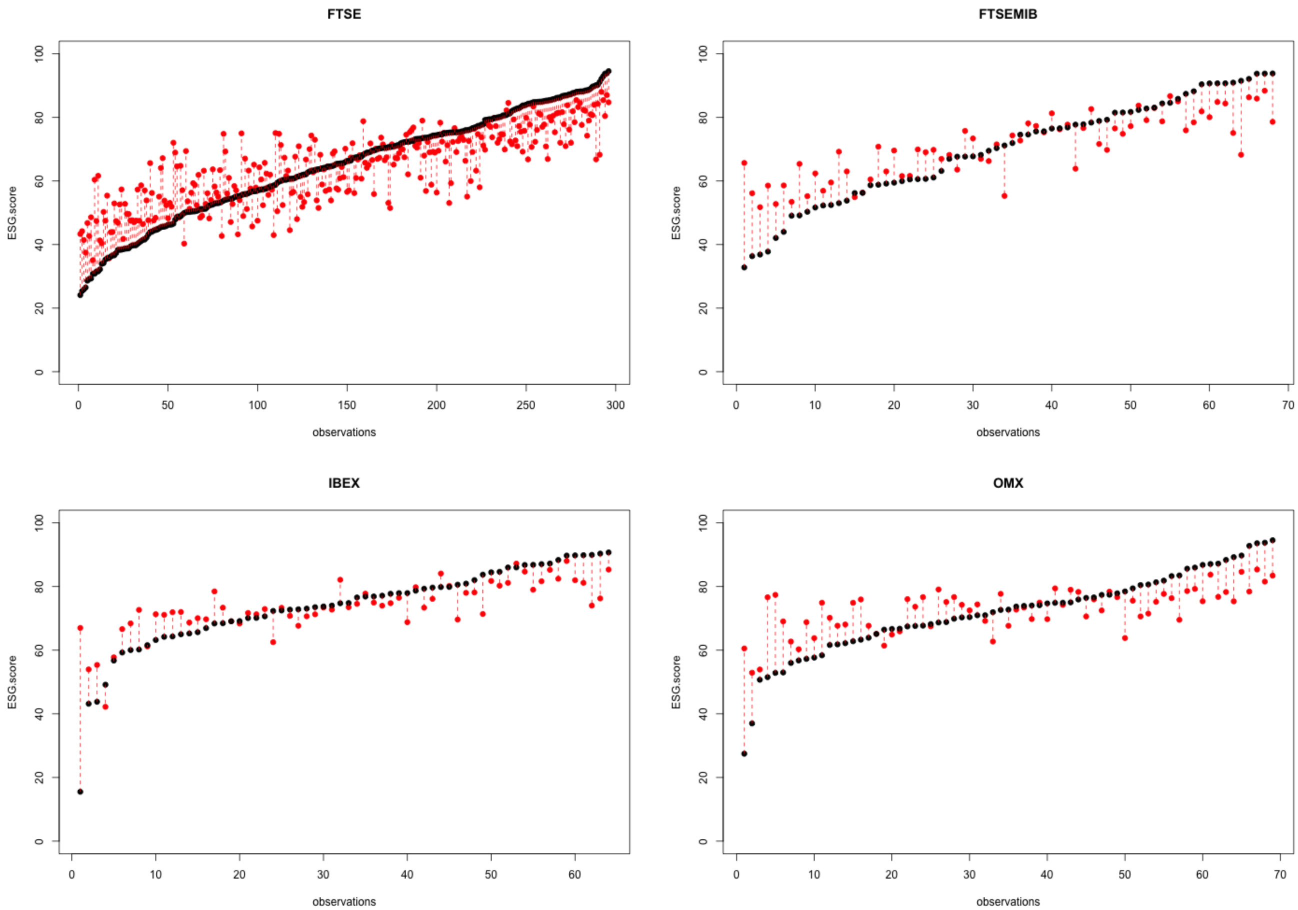


Figure 17: Predicted values (red dots) vs. observed values (black dots). Target variable: ESG score. Indexes: FTSE, MIB, IBEX, OMX.

5 Conclusion

The attention to ESG factors is becoming increasingly important to both regulators and investors. This is a natural consequence of the augmented demand for responsible and sustainable investments which mitigate risks associated with environmental and social issues, and improve the governance practices. As a result, companies are interested in understanding in which areas they should focus to improve their ESG features. This is a part of one of the most difficult challenge of the modern corporate strategy for companies across all industries: make profits through sustainable investments.

In this paper, we investigate the relationship between the ESG score and a list of accounting variables, or ratios, which summarize the company's activity and its wealth status. Also, we include the carbon intensity and the ESG controversial score as explanatory variables in the model to account for the non-financial characteristics of the firm. We assess whether the association mutates across different exchanges, to measure a potential country (policy) effect, and over the years, to study whether it changes alongside (i) different phases of the financial market, and (ii) variation in the attention to sustainability themes. We exploit the powerful random forest algorithm to detect also potential non-linear associations among variables, improving the results' interpretability by recurring to partial dependence plots and feature interaction.

We collect a sample of companies included into the eight major European stock indexes between 2012 and 2021 from Refinitiv. After a data cleaning process, we retrieve accounting, market, and sustainable data of 410 firms. We use the ESG score provided by Refinitiv as a proxy of the ESG features of the sampled companies. The largest part of the companies belongs to the principal European stock index (i.e., the FTSE, London Stock Exchange), which is also characterized by the largest volumes over the years.

A fist descriptive analysis reveals the ESG score positively related with its components and the Size of the firm, according to the Pearson correlation measure. Differently, it shows small, albeit negative, linear association with the ESG controversies score (ESGC), asset turnover (AT), ROA, and ROI. The indexes average ESG score exhibit an almost constant upward trend over the years, reflecting the results of the EU countries efforts to meet the UN sustainable development goals.

The main results on the ESG score prediction through the RF algorithm are as follows. The PDPs show the Size, approximated by the base 10 logarithm of the Total Assets, as the most important variable to explain the ESG score for several indexes (i.e., AEX, CAC, FTSE, MIB, and IBEX). Also, we observe a common behavior across them: the growth of the firms' size implies a non-linear increasing trend with jumps of the ESG Score. This result is in line with the current literature, which sees the largest companies characterized by the highest ESG ratings. A possible interpretation of this relationship relies on the fact that the largest firms are able to provide a better and more complete disclosure on ESG themes than the others, and they are rewarded accordingly.

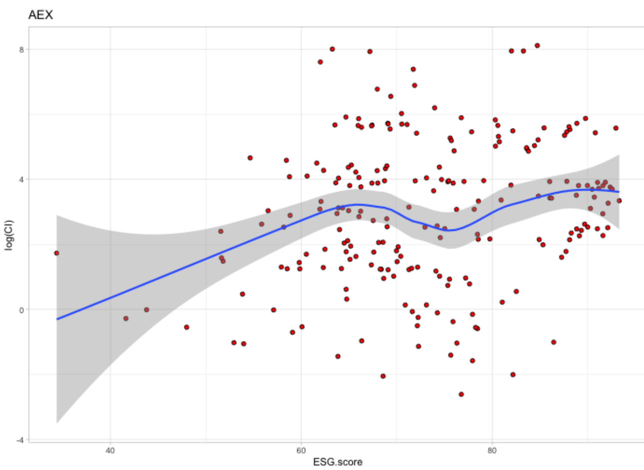
While the carbon intensity (CI) seemed uncorrelated with the other variables we considered, thanks to the machine learning ability to uncover complex non-linear relationship among features, it appears to be as an essential driver to explain the ESG score. The relationship is inverse: the higher the carbon intensity, the lower the ESG score. We attribute this evidence to the almost certainly significant association between the E pillar, which composes the ESG score, and the environmental ratio.

Another important variable according to the RF algorithm is the asset turnover (AT). For instance, the ESG score assumes the highest values when the AT ranges between 0.75 and 1.25 in the DAX and OMX indexes. However, when the AT is roughly 1.5, the target variable falls from 75 to 72 in the DAX index, and from 72.5 to 69 in the OMX index. Thus, it seems there is a negative non-linear relationship between these two variables.

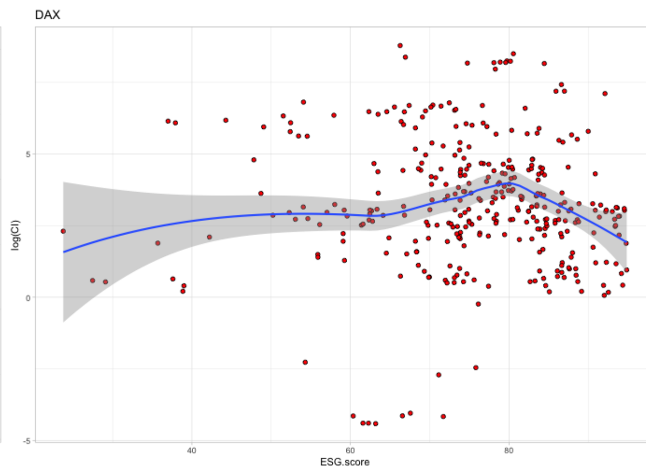
Interestingly, neither the year nor the business sector are associated to the ESG Score. However, we provide evidence of actual differences within the European stock exchanges in the way the explanatory variables impact the target variable. For instance, the smallest (according to the volumes) stock exchanges, namely the BEL and the OMX, include the ratio between the EBIT and the total revenues among the most important variables to explain the ESG Score. Differently, the largest ones often choose the firm's size.

This paper highlights the different importance of the market, accounting, and non-financial firms' features in predicting the ESG Score. It also reveals country, or policy-based discrepancies in predicting this sustainable rating. Nevertheless, this analysis emphasizes the strength of the machine learning methodology in recognizing, modeling, and interpreting potential non-linear relationships among variables.

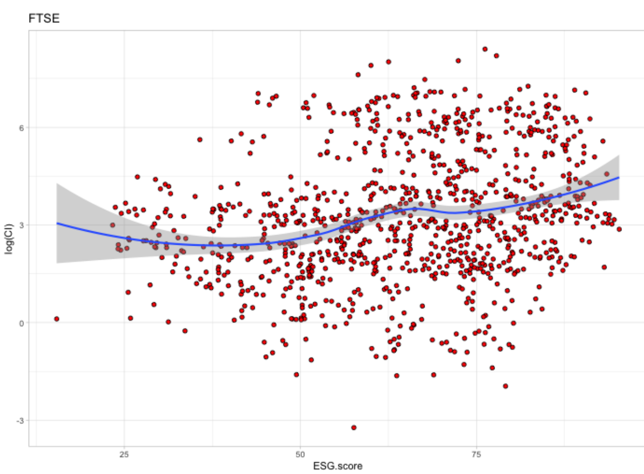
6 Appendix



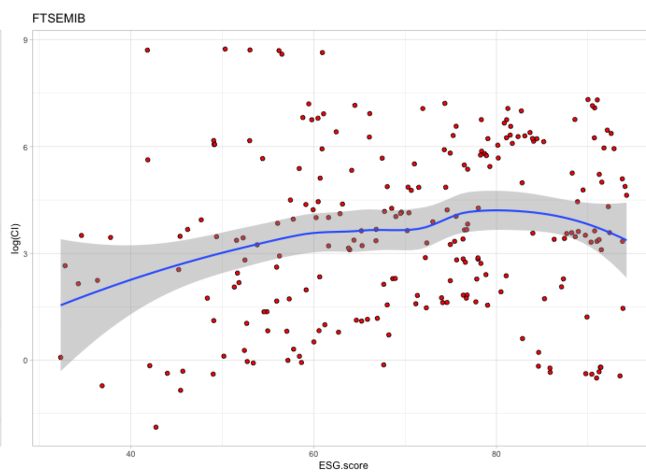
(a) AEX



(b) DAX



(c) FTSE



(d) MIB

Figure 18: Scatterplot of ESG score and $log(CI)$. Points represent the observed values, and lines represent locally estimated scatterplot smoothing (LOESS). Indexes: AEX, DAX, FTSE, MIB.

AEX							
Y	ESG	E	S	G	CI	Contr	Comb
2012	64.57	62.76	65.84	64.85	213.45	79.74	59.01
2013	65.48	63.54	67.57	64.77	274.14	82.40	61.21
2014	65.33	64.08	67.38	64.34	271.49	80.26	58.98
2015	67.25	66.66	70.66	64.52	361.49	91.17	64.30
2016	65.25	64.27	69.27	61.96	322.02	89.35	61.71
2017	68.39	66.08	73.47	63.50	271.59	87.91	64.21
2018	69.39	64.60	73.19	68.34	301.68	82.75	63.82
2019	70.15	67.10	75.20	66.15	285.30	81.53	64.63
2020	71.38	66.20	75.68	68.99	277.72	80.84	65.61
2021	74.60	69.27	75.69	77.57	93.37	85.45	69.95

BEL							
Y	ESG	E	S	G	CI	Contr	Comb
2012	44.50	46.28	40.40	50.35	134.66	91.27	43.45
2013	45.99	47.41	40.99	51.77	138.60	90.46	45.00
2014	45.40	45.45	42.04	50.26	136.84	91.29	43.99
2015	48.14	47.92	48.64	47.34	150.12	96.43	47.46
2016	49.68	48.61	52.72	47.10	143.83	93.40	48.85
2017	50.87	46.50	56.16	47.92	132.93	96.04	49.79
2018	57.17	53.64	60.98	55.63	106.74	93.43	55.44
2019	60.44	59.88	62.89	58.58	109.05	95.13	59.34
2020	62.58	63.81	66.07	57.24	99.12	93.02	61.55
2021	60.82	60.39	64.13	57.36	80.31	98.68	60.72

CAC							
Y	ESG	E	S	G	CI	Contr	Comb
2012	68.84	75.59	71.02	59.77	245.86	76.06	62.22
2013	68.70	75.48	71.43	59.00	229.58	76.92	62.17
2014	68.36	73.40	73.22	57.62	253.75	78.04	62.24
2015	70.45	75.78	77.15	56.97	291.07	87.53	67.24
2016	72.57	77.02	79.21	60.24	294.38	78.59	66.30
2017	74.44	76.92	81.69	62.10	243.26	81.85	69.09
2018	76.09	76.58	81.64	68.08	224.95	74.06	68.08
2019	76.85	78.76	82.82	66.77	207.66	72.08	68.04
2020	78.64	79.67	83.40	70.88	201.61	59.17	63.71
2021	79.51	78.95	83.66	74.69	201.06	62.72	66.61

DAX							
Y	ESG	E	S	G	CI	Contr	Comb
2012	69.07	71.97	72.87	61.42	311.43	70.33	60.02
2013	69.02	72.71	72.57	61.19	309.56	70.48	59.67
2014	69.16	73.25	73.11	60.18	337.10	74.90	62.03
2015	68.53	70.87	74.00	59.09	367.79	81.74	63.56
2016	69.86	71.87	76.74	58.92	391.74	67.92	60.88
2017	72.15	71.06	78.87	62.31	432.79	72.03	63.57
2018	74.35	71.28	78.66	69.94	421.07	74.46	66.97
2019	75.27	72.98	78.41	72.41	360.29	69.95	65.37
2020	79.00	75.21	80.15	80.33	301.18	64.73	66.25
2021	79.85	74.74	80.65	83.43	196.92	69.00	68.95

Table 3: Indexes average values per year of ESG, E, S, G, CI, ESGC, and ESG combined Score (Comb). Indexes: AEX, BEL, CAC, DAX.

FTSE

Y	ESG	E	S	G	CI	Contr	Comb
2012	57.26	58.93	56.31	58.07	145.58	77.82	51.62
2013	57.80	59.29	58.34	57.53	160.02	79.94	52.60
2014	58.14	58.85	59.53	57.43	156.98	79.20	52.95
2015	60.37	58.76	61.34	62.04	180.31	89.11	57.47
2016	61.67	59.74	62.90	62.90	183.53	84.76	57.64
2017	63.62	57.17	65.75	64.75	165.26	82.42	58.67
2018	64.33	57.15	66.79	65.35	153.61	83.01	59.80
2019	66.49	59.67	67.79	68.78	161.14	75.41	59.60
2020	69.35	61.55	69.61	73.75	185.14	71.56	60.73
2021	71.82	66.18	71.25	76.09	156.74	73.27	63.10

IBEX

Y	ESG	E	S	G	CI	Contr	Comb
2012	67.65	75.84	73.83	53.63	221.51	90.20	65.19
2013	67.62	75.22	74.14	54.10	227.44	93.55	66.53
2014	64.43	70.85	70.94	51.63	228.48	92.81	62.20
2015	67.71	72.30	72.86	57.66	257.72	95.07	66.35
2016	69.58	74.90	76.05	56.76	258.45	94.00	67.64
2017	70.63	74.13	77.04	57.69	229.17	91.79	68.75
2018	73.03	72.88	80.61	61.54	194.87	90.02	70.67
2019	73.93	75.02	82.09	61.50	165.33	83.44	69.72
2020	75.76	78.11	83.56	62.45	156.02	80.47	69.50
2021	76.32	80.29	82.72	63.47	137.17	72.95	67.51

MIB

Y	ESG	E	S	G	CI	Contr	Comb
2012	60.81	65.75	63.20	54.33	501.24	88.42	57.82
2013	62.49	67.47	63.82	57.46	215.94	86.35	58.85
2014	62.97	66.88	66.33	56.68	218.47	87.73	60.16
2015	60.73	65.75	64.44	53.90	249.77	93.52	59.05
2016	62.76	68.03	68.22	53.56	243.69	81.23	57.72
2017	64.43	63.20	69.38	57.18	373.62	89.83	61.10
2018	67.74	64.40	73.69	60.73	327.82	86.55	63.34
2019	70.61	69.52	77.26	62.19	316.53	81.51	64.47
2020	74.92	71.27	79.49	71.72	280.07	66.17	63.76
2021	77.68	74.12	82.34	74.78	363.27	80.00	69.03

OMX

Y	ESG	E	S	G	CI	Contr	Comb
2012	65.99	66.56	69.23	62.93	101.22	89.34	63.71
2013	65.85	67.22	69.76	61.25	111.43	84.34	62.28
2014	64.98	64.29	71.85	58.64	100.15	87.06	61.47
2015	66.10	63.40	73.62	61.06	95.03	91.74	64.52
2016	67.51	64.94	76.47	60.56	97.18	84.28	64.53
2017	69.80	66.52	80.14	60.26	75.99	84.03	66.51
2018	70.00	64.65	77.80	66.27	77.35	79.81	65.10
2019	69.47	64.72	74.98	68.45	77.16	79.68	64.13
2020	72.86	67.37	76.39	74.41	65.24	83.38	67.81
2021	74.51	71.39	78.74	72.53	60.11	80.15	68.79

Table 4: Indexes average values per year of ESG, E, S, G, CI, ESGC, and ESG combined Score (Comb). Indexes: FTSE, IBEX, MIB, OMX.

References

BBC (2015). Volkswagen: The scandal explained. <<bbc.com/news/business-34324772>>.

Berg, F., Koelbel, J. F., and Rigobon, R. (2022). Aggregate confusion: The divergence of ESG ratings. *Review of Finance*, 26(6):1315–1344.

Biau, G. and Scornet, E. (2016). A random forest guided tour. *TEST*, 25(2):197–227.

Bloomberg (2021). Deutsche Bank's DWS Slumps After U.S., Germany ESG Probe. <<bloomberg.com/news/articles/2021-08-26/dws-shares-fall-after-u-s-opens-probe-on-sustainability-claims#xj4y7vzkg>>.

Breiman, L. (2001). Random forests. *Machine Learning*, 45:5–32.

Breiman, L. (2003). *Manual - setting up, using and understanding random forests V4.0*. <<stat.berkeley.edu/~breiman/Using_random_forests_v4.0.pdf>>.

Clark, G. L., Feiner, A., and Viehs, M. (2015). From the stockholder to the stakeholder: How sustainability can drive financial outperformance. *Available at SSRN 2508281*.

D'Amato, V., D'Ecclesia, R., and Levantesi, S. (2021). Fundamental ratios as predictors of esg scores: A machine learning approach. *Decisions in Economics and Finance*, 44:1087–1110.

D'Amato, V., D'Ecclesia, R. L., and Levantesi, S. (2023). Firms' profitability and esg score: A machine learning approach. *Applied Stochastic Models in Business and Industry*, pages 1–19.

Eccles, R. G., Serafeim, G., Seth, D., and Ming, C. C. Y. (2013). The performance frontier: Innovating for a sustainable strategy: Interaction. *Harvard Business Review*, 91(7):17–18.

European Commission (2017). Dieselgate. <<bbc.com/news/business-34324772>>.

European Commission (2023a). Corporate Sustainability Reporting. <<finance.ec.europa.eu/capital-markets-union-and-financial-markets/company-reporting-and-auditing/company-reporting/corporate-sustainability-reporting_en>>.

European Commission (2023b). Sustainable Finance Disclosures Regulation. <<finance.ec.europa.eu/regulation-and-supervision/financial-services-legislation/implementing-and-delegated-acts/sustainable-finance-disclosures-regulation_en#:~:text=The%20Sustainable%20Finance%20Disclosures%20Regulation,laid%20down%20in%20the%20directive.>>.

European Union (2014). Directive 2014/95/EU of the European Parliament and of the Counsil. <<eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:32014L0095>>.

Financial Times (2022). German police raid DWS and Deutsche Bank over greenwashing allegations.

Flammer, C. (2015). Does corporate social responsibility lead to superior financial performance? a regression discontinuity approach. *Management Science*, 61(11):2549–2568.

Friede, G., Busch, T., and Bassen, A. (2015). ESG and financial performance: aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment*, 5(4):210–233.

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, pages 1189–1232.

Hartzmark, S. M. and Sussman, A. B. (2019). Do investors value sustainability? a natural experiment examining ranking and fund flows. *The Journal of Finance*, 74(6):2789–2837.

Jo, H. and Na, H. (2012). Does CSR reduce firm risk? Evidence from controversial industry sectors. *Journal of Business Ethics*, 110:441–456.

Jones Day (2021). The Impact of French and German ESG Due Diligence Laws on Business. <<jonesday.com/en/insights/2021/08/the-impact-of-french-and-german-esg-due-diligence-laws-on-business>>.

Kotsantonis, S. and Serafeim, G. (2019). Four things no one will tell you about ESG data. *Journal of Applied Corporate Finance*, 31(2):50–58.

Louppe, G., Wehenkel, L., Sutera, A., and Geurts, P. (2013). Understanding variable importances in forests of randomized trees. In Burges, C., Bottou, L., Welling, M., Ghahramani, Z., and Weinberger, K., editors, *Advances in Neural Information Processing Systems*, volume 26. Curran Associates, Inc.

Lourenço, I. C., Branco, M. C., Curto, J. D., and Eugénio, T. (2012). How does the market value corporate sustainability performance? *Journal of Business Ethics*, 108:417–428.

Morgan Stanley (2017). Sustainability Report. <<morganstanley.com/content/dam/msdotcom/about-us/giving-back/sustainability-at-morgan-stanley/2017_MS_Sustainability_Report.pdf>>.

Morningstar (2020). European Sustainable Funds Landscape. <<morningstar.com/en-uk/lp/sustainable-funds-landscape>>.

Oikonomou, I., Brooks, C., and Pavelin, S. (2012). The impact of corporate social performance on financial risk and utility: A longitudinal analysis. *Financial Management*, 41(2):483–515.

Popescu, I.-S., Hitaj, C., and Benetto, E. (2021). Measuring the sustainability of investment funds: A critical review of methods and frameworks in sustainable finance. *Journal of Cleaner Production*, 314:128016.

Refinitiv (2022a). Environmental, Social, and Governance Scores From Refinitiv. <<https://www.refinitiv.com/content/dam/marketing/en_us/documents/methodology/refinitiv-esg-scores-methodology.pdf>>.

Refinitiv (2022b). Environmental, Social and Governance Scores from Refinitiv. <<refinitiv.com/content/dam/marketing/en_us/documents/methodology/refinitiv-esg-scores-methodology.pdf>>.

Robeco (2023). ESG Funds. <<robeco.com/en-int/glossary/sustainable-investing/esg-funds>>.

SDSN Europe (2022). Europe Sustainable Development Report 2022. <<eu-dashboards.sdgindex.org>>.

United Nations (2012). Global Sustainable Development Goals. <<sdgs.un.org/goals>>.

World Economic Forum (2020). Billions for sustainable investments - Germany's plan for a green recovery. <<weforum.org/agenda/2020/07/germany-green-recovery-billions-sustainable-investments>>.

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

Ye, M., Wang, H., and Lu, W. (2021). Opening the "black box" between corporate social responsibility and financial performance: From a critical review on moderators and mediators to an integrated framework. *Journal of Cleaner Production*, 313:127919.

Zeb (2022). European Sustainable Investment Funds Study. <<zeb-consulting.com/en-DE/publications/european-sustainable-investment-funds-study-2022>>.