Vrije Universiteit Amsterdam KPMG

Master Thesis

Balancing Sustainability and Profitability: Quantifying Trade-offs Between ESG Compliance and Profitability in Financial Decision-Making

Author: Famke Bergsma (2770279)

1st supervisor: Daniele Avitabile
daily supervisor: Lindi Slabbert (KPMG, Financial Risk Management)
2nd reader: René Bekker

July 29, 2025

Abstract

Environmental, Social, and Governance (ESG) principles are increasingly used to guide corporate strategies and investment decisions. However, the relationship between ESG compliance and financial performance remains underexplored. Conducted in collaboration with KPMG, this research supports the firm's ESG advisory efforts by providing data-driven insights that promote sustainable and profitable decision-making. The research question of this study is: *How can the trade-offs between ESG compliance and profitability be quantified in financial decision-making?*

To address this, this study investigates the relationship between ESG performance and financial outcomes, and proposes a model to assess and optimize the trade-offs between sustainability and profitability. Generalized Additive Models (GAMs) are employed to capture the nonlinear and complex interactions between ESG variables and financial ratios. The analysis is conducted in both directions, predicting ESG performance based on financial metrics and vice versa. Two datasets are used, containing firms listed in Europe and the United States, and collected from the LSEG platform. Financial performance is measured using the following financial ratios: Earnings per Share (`EPS`), Return on Assets (`ROA`), and Return on Equity (`ROE`). ESG performance is assessed using variables including the `ESG Score`, `Environmental Score`, `Social Score`, `Governance Score`, and `ESG Combined Score`.

The simplest GAMs exhibited limited explanatory power, with R^2-values below 0.11 for both the European and USA datasets. However, when categorical control variables—`Market Capitalization`, `NAICS National Industry Name`, and `Country of Exchange`—were incorporated, the extended GAMs showed substantial improvements in predictive accuracy, achieving R^2-values between 0.4 and 0.65 for ESG prediction. In contrast, models that predicted financial ratios from ESG scores performed slightly worse, with R^2-values ranging

from 0.23 to 0.66. This suggests that financial performance is more difficult to predict.

Based on the best-performing GAMs, trade-off curves are plotted to visualize the interactions between ESG and financial performance. The results reveal clear directional patterns. In both Europe and the USA, higher `EPS` is consistently associated with stronger ESG scores, particularly along the `Environmental Score` and `Social Score`. However, `ROA` and `ROE` often show negative or nonlinear relationships with ESG scores, suggesting potential efficiency trade-offs. When reversing the direction of analysis, strong `Environmental Score` and `Social Score` are again linked to higher `EPS`, especially in the USA, highlighting asymmetries and regional differences. However, higher ESG is associated with a lower `ROA` and a lower `ROE`. Meaning that there is a negative relationship between ESG and `ROA` and `ROE`.

To identify optimal trade-offs, the Non-dominated Sorting Genetic Algorithm II (NSGA-II) is applied to derive Pareto-efficient GAM configurations. The final solution is selected from the Pareto front's Knee Point, representing the most balanced compromise between sustainability and profitability. In the European dataset, this Knee Point combines high `Environmental Score` and `Social Score` with strong `EPS`, suggesting that ESG integration can enhance profitability. In contrast, the American Knee Point emphasizes very strong `Environmental Score` and `Social Score` paired with a high `ROE`, while other metrics remain more moderate. These findings provide a grounded, quantitative basis for identifying ESG strategies aligned with financial objectives, tailored to regional and firm-specific contexts.

Contents

1

Introduction

In this chapter the introduction of this research will be discussed. In Section 1.1 the context of the research will be addressed. The problem will be explained in Section 1.2, and the relevance of this research will be explained in Section 1.3. The scope and limitations of this research will be addressed in Section 1.4. In the last Section of this chapter, Section 1.5, the thesis outline will be given to show how this thesis is structured.

1.1 Context of the Research

In recent years, the growing emphasis on Environmental, Social, and Governance (ESG) principles has fundamentally reshaped the financial landscape. ESG principles refer to a set of criteria used to evaluate a company's performance in areas beyond traditional financial metrics. These principles assess a company's impact on the environment, its relationships with employees, customers, and communities, and the strength of its leadership, ethics, and internal controls.

Since the 1970s, the theoretical discussion on sustainability and its financial implications for firms has evolved significantly. A pivotal moment in this debate was Nobel Laureate Milton Friedman's introduction of the Shareholder Theory, which argues that a manager's primary responsibility is to maximize shareholder value (1). This perspective sparked significant controversy in the emerging discourse on corporate sustainability, as it framed social and environmental responsibility as secondary to financial performance. Friedman's argument reinforced the notion that ESG considerations could only be justified if they contributed to profitability, shaping decades of debate on whether sustainability should be an ethical obligation or a strategic financial decision. Under this view, corporate social responsibility initiatives that do not enhance financial performance are seen as a deviation from a firm's primary objective.

However, in recent decades, alternative perspectives have gained prominence, challenging the traditional notion that financial success and sustainability are mutually exclusive. The Stakeholder Theory, introduced by Freeman (1984) (2), emphasizes that businesses should consider the interests of all stakeholders, including employees, customers, suppliers, and society at large, rather than focusing solely on shareholders. This shift has led to the rise of ESG frameworks, which aim to integrate sustainability into corporate decision-making while maintaining long-term financial performance.

To the best of my knowledge, D. Michael et al. (2019) published the first study examining both Shareholder Theory and Stakeholder Theory concurrently and identifying the predictor variables of each theory to assess their impact on corporate financial stability (3). Their findings suggest that companies adopting a stakeholder-oriented approach often exhibit higher resilience during financial downturns, as they benefit from strong relationships with employees, suppliers, and customers. This empirical evidence supports the argument that businesses considering ESG principles can better mitigate financial risks and enhance long-term profitability compared to those that do not.

The integration of ESG principles into financial decision-making represents a natural evolution of the Stakeholder Theory. The modern ESG framework emerged from growing social and environmental movements, including climate activism, corporate accountability campaigns, and investor-led sustainability initiatives (4). These movements reflected growing concerns from consumers and shareholders who were willing to sacrifice part of their financial interests to support enterprises committed to improving environmental, social, and governance aspects. Over time, ESG principles have evolved from an ethical consideration to a strategic necessity, influencing both corporate policies and investment strategies.

1.2 Problem statement

Research has shown that firms with strong ESG performance tend to experience lower capital costs, reduced regulatory risks, and improved brand loyalty, all of which contribute to financial stability (5). Companies integrating ESG principles into their business strategies often benefit from enhanced investor confidence and resilience during financial downturns.

However, ESG stocks tend to have lower short-term expected returns due to the additional constraints they must adhere to in order to meet ESG standards (6). Despite this, if

ESG performance improves market valuation and reduces financing costs, firms may have greater incentives to invest in sustainable business models and research and development, ultimately enhancing their ESG ratings (4). This raises the question of whether ESG compliance can be both a financial burden and a long-term value driver.

Given pressing global challenges such as climate change, resource scarcity, and evolving regulatory landscapes, understanding the financial trade-offs of ESG compliance has become increasingly important. Consequently, this study aims to examine the relationship between ESG and firms' financial performance by addressing the following research question: **How can the trade-offs between ESG compliance and profitability be quantified in financial decision-making?**

To structure the research, the problem will be divided into the following sub-questions:

- How can an ESG score be defined, and how can it be used to quantitatively measure ESG compliance and profitability across different firms and industries?
- What is the relationship between ESG scores and financial performance performance?
- How can a model be developed to analyze and optimize trade-offs between ESG compliance and profitability?

1.3 Relevance for KPMG

This research will be conducted in collaboration with the host organization, KPMG. KPMG is a global professional services firm offering audit, tax, and advisory services. The firm delivers innovative and strategic solutions to help organizations navigate complex challenges, such as mitigating climate change and promoting sustainable growth through technological advancements. KPMG's expertise spans multiple industries, including financial services, healthcare, and technology.

ESG is a top priority for KPMG, as the firm seeks to integrate sustainability into its business model and services. KPMG develops ESG solutions that provide clients with deeper insights while ensuring compliance with ESG laws and regulations. By offering advisory services that go beyond regulatory requirements, KPMG helps financial institutions integrate ESG metrics into risk assessment, investment strategies, and regulatory compliance frameworks.

This research aligns closely with KPMG's focus on sustainability, making it particularly relevant for the firm. The findings will contribute to the continuous improvement of KPMG's ESG solutions and advisory services, ensuring they align with both regulatory standards and the financial success of KPMG's clients by helping businesses optimize ESG integration without sacrificing profitability.

1.4 Scope and Limitations

This thesis aims to quantify the trade-offs between ESG compliance and profitability in financial decision-making. While ESG integration is often associated with risk mitigation, enhanced brand reputation, and long-term value generation, its short-term financial implications remain debated. Some firms may experience increased costs due to regulatory compliance, supply chain restructuring, and sustainability investments. Conversely, businesses that fail to adopt ESG principles may face reputational risks, regulatory penalties, and reduced investor confidence.

By examining financial data, industry trends, and empirical evidence, this research seeks to provide a comprehensive understanding of how ESG compliance affects profitability. Specifically, it will investigate whether companies that prioritize ESG principles achieve superior financial performance in the long run or if the associated costs outweigh the benefits. Understanding the financial dynamics of ESG compliance will enable firms to make data-driven decisions that balance sustainability with profitability, ultimately shaping the future of responsible investing.

The first limitation of this study is its geographic scope, as it includes only data from European and USA companies. As a result, firms from other regions of the world are not represented, which may limit the global generalizability of the findings.

Additionally, the analysis is restricted to variables for which data is readily available, specifically ESG scores and selected financial performance indicators. Due to time constraints, stemmingg from the six-month research collaboration with KPMG, the study focuses on twodata setsts (European and American) and does not incorporate a risk analysis. Financial performance is instead assessed using three key financial ratios.

These time constraints, combined with potential resource limitations, may have influenced the overall depth and scope of the research.

Moreover, the study depends heavily on the availability and quality of ESG and financial data. Since ESG reporting has only gained traction in recent years, the historical data is limited in coverage and duration. This may introduce bias or restrict the applicability of the results.

1.5 Organization

The structure of this thesis is as follows: Chapter 2 provides the background and literature review. Chapter 3 focuses on data analysis and preparation. The research models are introduced in Chapter 4, followed by their exploratory results and model development in Chapter 5. Chapter 6 presents and analyzes the results. Finally, Chapter 7 explores limitations, biases and potential directions for future research and Chapter 8 concludes the findings.

All of the code developed for this thesis is available on GitHub repository (Master Thesis).

2

Background and Literature Review

This chapter provides an overview of key concepts and existing research relevant to this study. Section 2.1 introduces ESG and its three pillars; Environmental, Social, and Governance. Section 2.2 discusses ESG compliance, and Section 2.3 outlines key financial performance indicators which assess a company's profitability and efficiency. Finally, Section 2.4 explores the link between ESG and financial performance.

2.1 ESG (Environmental, Social, Governance)

The ESG framework evaluates a company's Environmental, Social, and `Governance` performance, helping investors and stakeholders assess sustainability and ethical impact. While principles related to ESG have existed for decades, the term ESG was formally introduced in a 2004 United Nations report and has since been widely adopted across Europe, North America, and other developed economies (7).

ESG consists of three key areas:

- **Environmental (E)**: This aspect looks at how a company performs in terms of its impact on the planet. It includes factors such as carbon emissions, energy usage, waste management, and resource conservation (8). The environmental component of ESG has its roots in environmental awareness movements that gained prominence in the 1960s, largely influenced by works such as Rachel Carson's book *Silent Spring* (1962) (9). This seminal work documented the adverse effects of indiscriminate pesticide use and is widely credited with launching the modern environmental movement.
- **Social (S)**: This component focuses on a company's relationship with its employees, customers, suppliers, and the communities in which it operates. It encompasses

issues such as employee welfare, diversity and inclusion, labor practices, customer satisfaction, and community engagement (10).

- **Governance (G)**: Involves the leadership and management structure of the company. It looks at the transparency, accountability, and integrity of a company's decision-making processes. This includes board diversity, executive compensation, shareholder rights, and corporate ethics (11).

Firms receive ESG scores that indicate their level of ESG adherence. The higher the ESG score, the greater the compliance. These scores are typically measured on a scale from 0 to 100 (12), where 100 represents the highest possible ESG performance and 0 the lowest.

ESG scores are determined by agencies such as S&P, MSCI, Amundi, and Sustainalytics, each using distinct methodologies to evaluate a company's ESG performance. For instance, S&P applies a 0-100 scale, while MSCI uses a letter grading system ranging from AAA (highest) to CCC (lowest) to assess a company's relative ESG risk exposure and management effectiveness (13). In contrast, Amundi categorizes financial products based on their sustainability characteristics, using a scale from A (best) to G (worst) to enhance transparency in sustainable investments (14). Sustainalytics, on the other hand, measures ESG performance using a risk-based approach, assigning companies a score from 0 to 50, where 0 represents negligible ESG risk and 50 indicates the highest ESG performance (15).

Each company is assessed based on approximately 120 ESG-related questions, with a maximum of 1,000 question-level evaluations per company. These questions follow established scoring frameworks that evaluate four key aspects:

- **Availability:** Whether the company discloses sufficient ESG-related information publicly, such as sustainability reports, carbon emissions, or governance policies.

- **Quality:** The reliability and accuracy of the disclosed ESG data, ensuring it is comprehensive and free from inconsistencies.

- **Relevance:** How meaningful the reported ESG data is in assessing the company's actual sustainability performance, considering industry-specific factors.

- **Performance:** How well the company meets ESG criteria based on the disclosed information.

The scores from these individual questions are aggregated into criteria-level scores, which focus on the most material ESG themes relevant to a company's sub-industry. A company can have up to 30 criteria-level scores. These criteria-level scores are then grouped into the three main ESG categories: Environmental, Social, and Governance. Finally, these three category scores are combined into a single, overall ESG score for the company (12).

2.2 ESG Compliance

An ESG score remains a relatively subjective metric, raising the question: *What threshold defines a sufficiently high ESG score to qualify as compliant?* In cases where a firm receives multiple ratings from different agencies, the average of the scores may be used (12).

There is no universally agreed-upon threshold for ESG compliance, as it depends on the rating agency. S&P Global ESG Scores typically rate companies on a scale of 0 to 100, with a score above 70 considered strong, while anything below 50 may indicate poor ESG performance (12). For MSCI ESG Ratings, companies rated AA or AAA are considered ESG leaders, while those rated B or CCC are seen as having poor ESG practices (13). Similarly, Amundi's ESG rating system categorizes companies with an A rating as ESG leaders, while those rated B or C are considered to have poor ESG practices (14). For Sustainalytics, an ESG risk score below 10 indicates negligible risk, while scores above 40 signify severe ESG risks (15).

Globally, ESG reporting requirements remain inconsistent, with laws and regulations varying across jurisdictions (16). As noted, "*The ESG reporting landscape is dynamic, fragmented, and evolving.*" (16). This variation highlights the importance of recognizing that companies in different regions may face distinct drivers behind their ESG performance. Moreover, differing legal and regulatory frameworks can shape corporate ESG behavior in ways that diverge from those of companies in other countries.

2.3 Indicators of company's financial performance

Financial ratios are essential tools for evaluating a company's financial performance. These ratios facilitate the visualization and comparison of financial performance across multiple firms and different scenarios. In this study, they serve as key indicators for comparing a firm's financial performance with its ESG performance. The financial ratios considered in this research are detailed below.

Return on Assets (ROA) evaluates how effectively a company utilizes its assets to generate profit (17), as expressed in Equation 2.1.

$$\text{Return on Assets } (\texttt{ROA}) = \frac{\text{Net Profit}}{\text{Total Assets}} \times 100\%. \tag{2.1}$$

ROA combines profit margin and asset turnover, making it a crucial indicator of how well a company generates returns for investors and creditors (18). Research also highlights that firm value increases significantly with higher ROA (19).

Return on Equity (ROE) measures a company's ability to generate profit from shareholders' equity. A higher ROE indicates strong financial performance and efficient use of equity capital, while a lower or negative ROE may signal financial difficulties (18).

$$\text{Return on Equity } (\texttt{ROE}) = \frac{\text{Net Profit}}{\text{Shareholders' Equity}} \times 100\%. \tag{2.2}$$

ROE is a widely used metric by investors and corporate leaders to assess how effectively a company generates profit relative to its shareholders' equity. For investors, analyzing ROE is crucial as it helps evaluate the potential returns on their investments, while for companies, it serves as an attractive factor for potential investors (20).

Earnings per Share (EPS) quantifies the company's profitability per outstanding share of stock, providing a key metric for investors to assess financial performance (18, 21).

$$\text{Earnings per Share } (\texttt{EPS}) = \frac{\text{Net Profit}}{\text{Number of Outstanding Shares}}. \tag{2.3}$$

EPS is a widely used indicator of a company's profitability, with higher values signifying better financial health and being associated with improved performance and profitability for the firm (20).

2.4 Relationship between ESG and company's financial performance

The relationship between ESG performance and firm value has been widely examined in both academic and business research. Some studies suggest that ESG adoption can have adverse financial effects, as the costs of implementing sustainability initiatives may outweigh their financial benefits, potentially leading to lower profitability and weaker stock performance. This argument aligns with the traditional Shareholder Theory, by Friedman (1970) (1), which posits that a firm's primary responsibility is to maximize shareholder value.

Supporting this view, Di Giuli et al. (2014) (6) find that increases in ESG ratings are associated with negative stock returns and a reduced Return on Assets (ROA), suggesting that ESG investments may not always translate into immediate financial gains. Similarly, Lee et al. (2009) (22) and Filis et al. (2016) (23) report a negative relationship between ESG performance and financial outcomes when measured using market-based metrics and Return on Capital (ROC), respectively.

Brammer et al. (2006) (24) further contribute to this debate by showing that firms with low ESG ratings tend to outperform the market. However, their results lack statistical significance, raising questions about whether weak ESG performance might, in certain cases, lead to superior financial returns, particularly in industries where sustainability initiatives are costly or do not align with consumer preferences.

Offering a more nuanced perspective, Barnett et al. (2006) (25) identify a curvilinear relationship, where firms with moderate ESG adoption experience lower financial returns, while both weak and strong ESG integration are associated with higher performance. This suggests that firms must carefully balance their ESG commitments to optimize financial outcomes, as partial ESG adoption may lead to inefficiencies without delivering the full reputational or operational benefits of a well-integrated ESG strategy.

While some studies highlight a negative link between ESG and financial performance due to the costs associated with sustainability initiatives, a growing body of research suggests that ESG adoption can lead to positive financial outcomes. Evidence increasingly indicates that companies integrating ESG considerations into their strategies can achieve a balance between financial returns and social responsibility, reinforcing the notion that sustainability and profitability are not mutually exclusive.

One influential study by Bassen et al. (2015) (26) analyzes over 2,000 empirical studies conducted between the 1970s and 2015, examining the relationship between ESG criteria and corporate financial performance (CFP). Their findings reveal that approximately 90% of the studies report a nonnegative ESG–CFP relationship, with a consistently positive correlation observed since the mid-1990s. This extensive body of research provides strong evidence supporting the business case for ESG investing.

Another notable study by Atz et al. (2021) (5) examines over 1,000 studies conducted between 2015 and 2020, analyzing the relationship between ESG and financial performance. Their findings indicate that 58% of corporate-focused studies, those assessing operational metrics such as Return on Equity (ROE) and Return on Assets (ROA), report a positive relationship, while only 8% show a negative correlation (5). A positive relationship suggests that firms with stronger ESG performance tend to achieve higher ROE and ROA, potentially due to improved risk management, increased operational efficiency, and stronger investor confidence. These results highlight `ROE` and `ROA` as valuable indicators for measuring financial performance in the context of ESG.

Guo et al. (2018) (27) investigated the relationship between ESG and financial performance within the power sector. Their findings confirm a positive correlation, indicating that stronger ESG performance is associated with improved financial outcomes. As the authors state, "*The results show that good ESG performance can indeed improve financial performance, which has significant implications for investors, company management, decision-makers, and industry regulators.*" (27). This study reinforces the business case for ESG integration, suggesting that companies in the power sector can enhance financial performance by strengthening their ESG practices.

Research by Almeyda et al. (2019) (28), found a significant impact of ESG disclosure on the financial performance of real estate companies, as measured by accounting indicators such as `ROA` and Rate of Change (`ROC`). The study, which analyzed data from 77 listed real estate companies over a five-year period, revealed a statistically significant positive relationship between environmental disclosure and both firm `ROC` and stock price. However, it also concluded that the social and governance pillars had no significant influence on the financial performance of these firms (28).

A recent study from 2025 by Hong-Yi Chen et al. (29) investigates the impact of ESG factors on a firm's financial performance and risks using the SASB framework. This study used 1,544 firm-year observations with available ESG information and financial variables as data. The key findings indicate that traditional ESG metrics, such as the ESG Disclosure Score, do not effectively predict financial performance. However, the Sustainability Accounting Standards Board (SASB) ESG Score, which focuses on material ESG issues, significantly enhances financial performance, particularly through improved profit margins, market competitiveness, and operational efficiency. Additionally, the study highlights that a higher SASB ESG Score helps firms mitigate firm-specific risks and reduce stock price crash risk, suggesting an insurance-like effect of ESG investments. These findings imply that firms should prioritize material ESG issues to balance sustainability with profitability effectively.

Ahmed et al. (2025) (30) investigate how ESG performance influences firm value, particularly considering the role of cash holdings as a moderating factor. The study utilizes a comprehensive dataset encompassing 1,144 companies across 27 European Union countries over an 11-year period from 2013 to 2023. The data was sourced from LSEG Data & Analytics, providing a robust foundation for analysis. Employing panel regression techniques, specifically fixed and random effects models, the authors explore the relationship between ESG performance and firm value. The Hausman test indicated that the fixed effect model was more appropriate for this analysis. The findings reveal a positive and statistically significant impact of ESG performance on firm value, suggesting that companies with efficient and effective ESG practices tend to have higher market valuations. Additionally, the study finds that higher cash holdings positively affect market value, indicating that substantial cash reserves can enhance a company's valuation. The authors recommend that board members in the EU region adhere to ESG principles and maintain higher cash reserves to substantially increase their company's value. These results are robust, as corroborated by similar outcomes obtained using Tobit regression analysis.

On the one hand, empirical evidence suggests that strong ESG performance is associated with superior financial outcomes. Glushkove et al. (31) illustrate that firms with high ESG ratings deliver better returns compared to companies with lower ratings, suggesting that ESG integration can enhance corporate financial performance through improved risk management, operational efficiency, and stakeholder trust. Similarly, Barnett (32) acknowledges these benefits but highlights their variability, noting that the financial advantages of ESG adoption depend on stakeholders' ability to exert influence and the firm's industry context. This aligns with the Stakeholder Theory perspective by Freeman (1984) (2), which argues that firms that effectively manage relationships with key stakeholders, such as customers, employees, and regulators, can achieve long-term financial stability and competitive advantage.

Gülay et al. (33) shows that there is a positive and highly significant relationship between ESG performance and firm value (coefficient: 0.008) as well as profitability (coefficient: 0.049). These results support the case for corporate managers to allocate more resources toward ESG initiatives. Additionally, they provide valuable insights for policymakers in developing measures that further promote ESG adoption.

Chia et al. (34) constructed a few machine learning models to examine the relationship between ESG scores and firm's `ROE`. For this study they used data sourced from Thomson Reuters DataStream where they used ESG scoring criterion together with the financial performance of selected companies over a period of five years. Chia et al. developed models like Support Vector Machine (SVM), Random Forest, Naive Bayesian, Multilayer Perceptron (MLP) Neural Networks, and Long Short-Term Memory (LSTM) Neural Networks. In the results, they showed that the Neural Networks based models generally achieved the highest accuracy of 81.08% of predictability based on MLP model. This accuracy is evaluated using metrics like mean squared error, root mean squared error, R^2-score for regression, and percentage of accuracy for classification.

Overall, the literature presents mixed findings, indicating that the financial implications of ESG adoption are complex and highly context, dependent. While strong ESG performance can enhance financial stability and investor confidence, particularly in industries where sustainability serves as a competitive advantage, the associated costs and constraints may also hinder short-term profitability. These contrasting perspectives underscore the need for further empirical research to determine the conditions under which ESG investments create or erode shareholder value.

A key challenge in analyzing the financial impact of ESG is the inconsistency in how companies report ESG data. Variations in disclosure standards and methodologies make it difficult to compare ESG performance across firms and industries, complicating efforts to assess its true financial effects (35).

3

Data Preparation

In this chapter, the data used in this study will be described, and the data preparation process prior to modeling will be outlined. In Section 3.1 the source of the data and the data will be described. Subsequently, Section 3.2 outlines the data cleaning process, which includes assessing data availability, visualizing the data, validity check, applying outlier detection, and handling missing values.

3.1 Data description

The data for this research is sourced from the London Stock Exchange Group (LSEG) database, previously known as Refinitiv. LSEG is well-known for its extensive coverage of ESG data, offering insights into sustainability metrics alongside financial indicators. The LSEG ESG database houses ESG scores for public and private companies globally, assessing and benchmarking their performance on key sustainability metrics. These scores rely on publicly reported data to evaluate a company's ESG performance. LSEG ESG data has been widely used in academic and professional research, serving as a reliable source for analyzing the relationship between ESG performance and financial outcomes (36).

From LSEG two different datasets are received. The first one is *Ossiam Stoxx Europe 600 ESG Equal Weight NR UCITS ETF (ETF).* This ETF (Exchange-Traded Fund) aims to replicate the performance of the STOXX Europe 600 ESG Broad Market Equal Weight Index, which tracks European companies. All included stocks are equally weighted (37). This dataset contains data of 479 firms that are all established in Europe. The other dataset is *SPDR S&P 500 ESG ETF (EFIV).* This ETF aims to track the performance of the S&P 500 Scored & Screened Index before fees and expenses. This index is designed to include S&P 500 companies that meet specific ESG criteria, while maintaining industry group weights similar to those in the S&P 500 Index (38). This dataset contains data of 315 firms that are all established in the USA.

3.1.1 Variables

The dataset includes company-specific information, ESG metrics, and financial ratios. The company-specific information for each firm includes the following variables:

- **Company Name**: The name of the company.
- **ISIN Code**: An ISIN (International Securities Identification Number) is a 12-character alphanumeric code used to identify a specific security, such as a company's stock or bond. It helps uniquely identify securities across different countries.
- **Ticker Symbol**: Is a unique series of letters assigned to a publicly traded company's stock for trading on a stock exchange. Like the ISIN Code, it serves as an identifier for a particular security.

Since the dataset contains ESG and financial information from various companies, several control variables are also included. The following three control variables are part of the dataset:

- **Country of Exchange**: Indicates the country in which the company is established. This variable is important to account for differences in ESG reporting standards across countries. For instance, the European Union has stricter ESG regulations compared to the United States. The European dataset includes companies from 17 different countries (see Table 8.1), whereas the USA dataset contains companies from only one country: the United States.
- **NAICS National Industry Name**: Refers to the industry classification assigned to a company based on the North American Industry Classification System (NAICS). NAICS is a standardized system used in the USA, Canada, and Mexico to classify businesses according to their primary economic activity. Since ESG risks and opportunities vary by industry, this classification is important. For example, a technology company and an oil company are evaluated on different ESG factors.

  The European dataset includes 203 different industries; the ten most frequent ones are listed in Table 8.2(a). The USA dataset includes 155 different industries, with the ten most frequent shown in Table 8.2(b).

- **Market Capitalization**: Refers to the total value of a company's outstanding shares of stock and is commonly used as a proxy for company size. Larger firms generally have higher `Market Capitalization`, while smaller firms have lower `Market Capitalization`.

These three control variables are selected based on the Moody's article *ESG Score Predictor: A Quantitative Approach for Expanding Company Coverage* (39). According to this article, company location, size, and industry are the three primary drivers of ESG performance. Among these, company size emerges as the most influential factor in determining a firm's ability and willingness to implement sustainable business practices.

The dataset contains the following ESG information per company per year:

- **ESG Score**: The general ESG score provides an overall assessment of a company's performance across `Environmental`, `Social`, and `Governance` factors.

- **Environmental Pillar Score**: This score evaluates the first aspect of the ESG concept, namely a company's environmental impact, policies, and initiatives.

- **Social Pillar Score**: This score assesses a company's social responsibility, including employee welfare, diversity and inclusion, human rights policies, and community engagement.

- **Governance Pillar Score**: This score reflects the quality of a company's corporate governance structure. It considers aspects such as board diversity, executive compensation, shareholder rights, business ethics, and transparency in decision-making.

- **ESG Combined Score**: The combined score integrates the overall ESG score while adjusting for significant controversies a company may face. It penalizes firms involved in major ESG-related incidents, providing a risk-adjusted measure of ESG performance.

These ESG variables are all scaled from 0 to 100, with 0 being the lowest and 100 the highest score (36). For a further explanation of the ESG variables, see Section 2.1.

Finally, the dataset includes the following financial ratios: **Return on Equity (`ROE`)**, **Return on Assets (`ROA`)**, and **Earnings per Share (`EPS`)**. For a detailed explanation of these financial ratios, see Section 2.3.

3.2 Data cleaning

Data cleaning is a crucial step in preparing the dataset for analysis, as it ensures the accuracy, consistency, and integrity of the data, thereby reducing the risk of biased or misleading results in subsequent statistical and empirical evaluations. In this study, the data cleaning process consist of three main stages: checking for duplicates, handling missing values, and detecting outliers.

To ensure data integrity and traceability, it was first verified that all company records were unique and associated with identifiable company information. Specifically, the columns `Company Name`, `ISIN Code`, and `Ticker Symbol` were examined for missing values and duplicate entries. No missing values or duplicate entries were found, indicating that these identifiers are complete and consistent across the dataset.

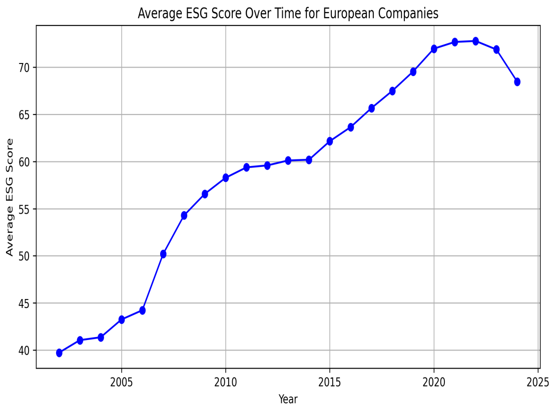
This results in a European dataset and a USA dataset, comprising information from a total of 794 firms, 479 in the European dataset and 315 in the USA dataset. The data obtained from LSEG was originally structured such that, for each ESG variable, financial ratio, and `Market Capitalization`, the dataset contained 25 columns listing the corresponding dates, followed by 25 columns with the associated values for each date. To clarify this format, a brief overview of five financial years is provided in Table 8.3.

To facilitate effective analysis and comparison across firms and years, the data needed to be reshaped so that each observation corresponds to a unique company-year combination. The first step involved determining the value of each variable for every company in each year. To accomplish this, a Python script was developed to extract the relevant yearly values and organize them into a structured DataFrame, as illustrated in Table 8.4. Subsequently, the data was transformed from wide format to long format, ensuring chronological sorting of observations for each company based on the `Year` variable (see Table 8.5 for an example). This long format enables easier temporal comparison of ESG scores and financial ratios, and provides a clear organization of the data across time for each company.

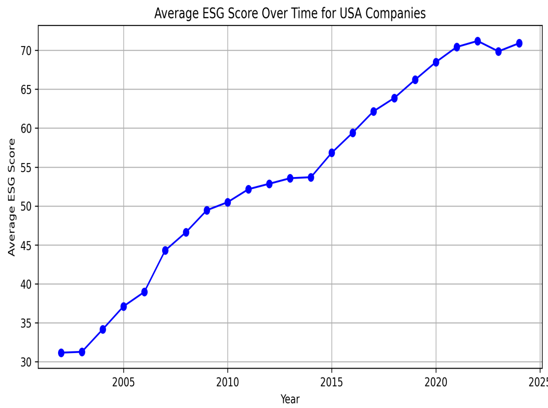
In the wide format, the European dataset consisted of 479 rows (each representing a unique company), with multiple columns for each year's data. After transformation, the long format resulted in 9,580 rows for the European data. Similarly, the USA dataset was transformed from 315 rows in wide format to 6,300 rows in long format.

3.2.1 Data Availability

For each firm, data is collected over a 25-year period. However, not all firms begin documenting their ESG and financial information in the same year. As a result, there are



(a)



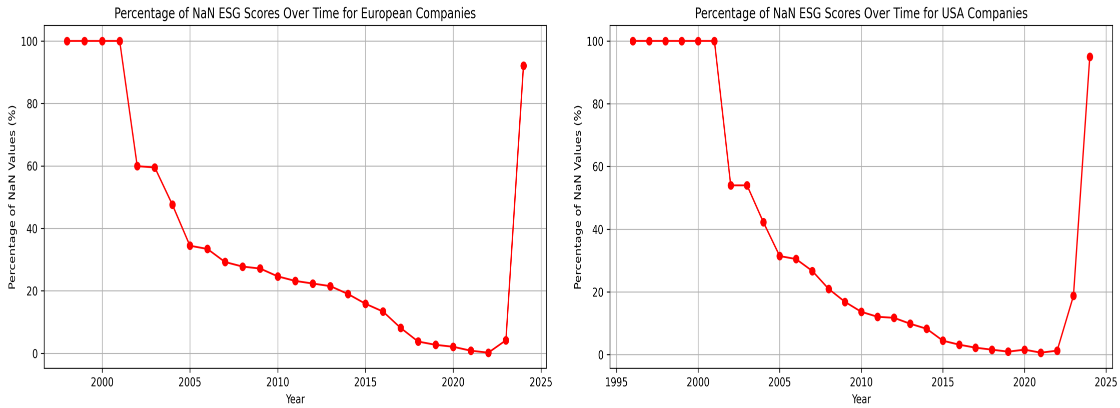
(b)

Figure 3.1: Average ESG score per year for (a) European Companies and (b) USA companies.

missing values (`NaNs`) for years when specific data is unavailable for a firm. For the European data, LSEG first ESG reported year is 1998 and the last is 2024, while financial ratios are reported from 2000 to 2025. The control variable, `Market Capitalization`, is reported for the years 1999 to 2025. For the USA data, LSEG is starting with reporting the ESG information in 1996 and last reporting in 2024, financial ratios from 2000 to 2025, and the control variable, `Market Capitalization`, has the first reported value in 2000 and the last in 2025.

To decide which time period is appropriate for this study, the average ESG score and the percentage of missing ESG values will be plotted per year, separately for the European and USA datasets. A suitable time period is one where the percentage of missing ESG data is relatively low, meaning that for most companies, ESG scores are available and consistently reported. At the same time, there will be looked for a period in which the average ESG scores do not fluctuate dramatically year-over-year, which would suggest that the scoring methodology is stable and the data is more reliable.

The average ESG score per time for European and USA data are plotted in Figure 3.1(a) and 3.1(b) respectively. Both datasets follow a similar trend. ESG scores begin relatively low and increase over time, with a notable rise from 2006 onward. European companies surpass an average score of 70 shortly after 2019, whereas USA companies reach this threshold in 2021. These elevated scores continue until 2023. The slight drop observed in 2024 may reflect reporting delays or incomplete data rather than an actual decline in ESG performance.



(a) (b)

Figure 3.2: Percentage of missing values in ESG scores per year for (a) European Companies and (b) USA companies.

The percentage of missing values of the ESG scores for the European and USA companies are plotted in Figure 3.2(a) and 3.2(b) respectively. Figures 3.2(a) and 3.2(b) show a high percentage of missing values in the early years. For the European dataset, 100% of ESG data is missing until 2001, and for the USA dataset until 2002. From 2003 onward, the availability of ESG scores improves steadily. In the European dataset, missing values drop to nearly 0% by 2022 but rise again in 2024. The USA dataset follows a similar trend, though a small portion of missing values persists even after 2022, with a notable increase again in 2024.

Due to the high percentage of missing values in 2024, this year will be excluded from the analysis. From 2004 onward, both regions show more reliable and complete ESG data coverage. Financial year 2024 is showing a lot of missing values and a decline in average ESG score. Therefore, the period from 2004 to 2023 is selected for this study.

3.2.2 Data Visualization

Before proceeding with data cleaning, including detecting outliers and handling missing values, the datasets were initially explored through visual inspection. Specifically, scatter plots were created for each variable over time for both the European and American datasets. These visualizations are included in the Appendix as Figure 8.1 and Figure 8.2.

Notably, the financial ratios exhibit considerable outliers and a broad range of variability across years in both datasets. This indicates that while the majority of firms cluster around a stable central tendency, some firms report extreme values. For instance, `EPS` values in the European dataset display extreme outliers both above 750 and around –750, whereas the American dataset only shows extreme negative outliers near –400. A similar pattern is observed for `ROA`, with an outlier around 500 in the European dataset and one near –500 in the American dataset. In contrast, `ROE` exhibits a wide spread in both datasets, without a dominant pattern, indicating variability across firms and time.

The ESG variables present a more structured and consistent pattern. In both datasets, ESG scores become denser and more complete in later years, particularly after 2015. This trend reflects the growing emphasis on ESG reporting and its broader adoption over time, aligning with global regulatory and investor, driven shifts toward sustainable finance.

`Market Capitalization` displays a right-skewed distribution that increases over time, consistent with the expansion of firm size and valuation in recent years. The American dataset shows denser observations and generally higher `Market Capitalization` values compared to the European dataset, suggesting differences in firm scale between the regions.

Across all scatter plots, some noticeable gaps and irregularities are present, especially in earlier years. These suggest missing or inconsistently reported data. To address these issues and ensure data quality, the next steps involve treating missing values and detecting outliers as part of the data cleaning process.

3.2.3 Validity Check

After inspecting the data availability, the next step was to verify whether all variables fell within their valid value ranges. This was done by visualizing and inspecting the minimum and maximum values for each variable. For the ESG variables, scores are expected to range between 0 and 100 (see Section 2.1). Upon inspection, all ESG values in both datasets fell within this valid range. Therefore, no apparently invalid values were detected in the ESG variables based on this check.

For the financial ratios there are no strict universal bounds. These variables can take on highly negative or positive values due to accounting distortions (e.g., extremely low equity) or extraordinary events.

For the `Market Capitalization`, values are expected to be non-negative, as negative `Market Capitalization` is not meaningful. A review confirmed that all `Market Capitalization` values in both datasets were greater than zero, and thus no invalid entries were found.

After verifying that all ESG values fall within valid ranges, the next step is to assess whether there are unreal fluctuations in the ESG variables, as these may indicate potential reporting errors or inconsistencies. In this study, an unreal fluctuation is defined as a year-over-year change of more than 15 points (either upward or downward) in an ESG score. This threshold is based on the Refinitiv ESG methodology (2023) (40) and supported by empirical findings in Berg et al. (2022) (41), which highlight that typical ESG score changes are gradual, and large deviations often stem from methodological revisions or data anomalies.

Rows exhibiting unreal fluctuations across all ESG variables were first identified. These rows were then analyzed to assess whether the fluctuations were concentrated within a limited number of companies. If that had been the case, the corresponding companies would have been excluded entirely from the dataset, as such concentrated anomalies may indicate systematically unreliable ESG reporting. However, this was not the case, as the anomalies were distributed across a wide range of companies, 44 in the European dataset and 45 in the USA dataset, rather than being concentrated in a select few. Therefore, only the specific rows displaying simultaneous unreal fluctuations in all ESG variables were removed. This process resulted in the exclusion of 48 rows from the European dataset and 51 rows from the American dataset.

3.2.4 Outlier Detection

After the validity check, the dataset was examined for outliers, as these can significantly affect the accuracy and reliability of the analysis. An outlier is a data point that deviates significantly from other observations in a dataset (42). As defined by Hawkins (1980) (43), *"An outlier is an observation that deviates so much from other observations as to arouse suspicions that it was generated by a different mechanism."* In the context of ESG and financial data, outliers may reflect errors, unusual events, or true but rare phenomena.

When dealing with panel data, such as multiple companies observed over time, as in this study, it is important to account for entity-specific dynamics. Each firm may differ in scale, reporting behavior, and industry characteristics, making global outlier detection (across all firms) potentially misleading. Therefore, this study applies outlier detection at the company level, identifying observations that are extreme relative to a company's own historical distribution.

The Interquartile Range (IQR) method is chosen for this purpose. It is a robust statistical technique that identifies outliers based on the spread of the middle 50% of data (44). For each company and for each variable of interest, the 25th percentile (Q1) and the 75th percentile (Q3) are calculated, and the IQR is defined as:

$$\mathrm{IQR} = Q_3 - Q_1.$$

Outlier thresholds are then set using a multiplier k, such that:

$$\text{Lower bound} = Q_1 - k \cdot \mathrm{IQR},$$
$$\text{Upper bound} = Q_3 + k \cdot \mathrm{IQR}.$$

Any value outside this range is flagged as an outlier. While the traditional IQR method uses $k = 1.5$, this study adopts a more lenient threshold of $k = 3.0$ to reduce the risk of false positives. This choice is motivated by the inherent variability in financial and ESG indicators (described in Section 3.2.2), as well as by the limited number of observations for certain firms. Using $k = 3.0$ focuses the detection on more extreme deviations while retaining the majority of valid data points (42).

The IQR method was applied to the ESG variables, financial ratios and the `Market Capitalization`. The detection was implemented using the `groupby` function in `pandas`, iterating over each firm to compute the IQR and flag company-specific outliers accordingly.

This approach aligns with the notion of contextual outliers, as emphasized by Aggarwal (2017) (45), who states *"Contextual outliers are defined relative to specific contextual attributes, which influence the expected behavior. For example, in time-series or panel data, an observation may be an outlier with respect to its own history rather than the overall data distribution."*

After detecting outliers, rows containing at least one flagged variable were removed from the dataset. While deleting outliers can result in information loss, the number of such cases is limited in this study. Therefore, full removal was deemed acceptable to preserve data quality for subsequent analyses (46).

As a result, 770 rows were removed from the European dataset, leaving 8,762 rows. Similarly, 509 rows were removed from the American dataset, leaving 5,740 rows.

3.2.5 Missing Values

Lastly, and before conducting any analysis, it is essential to assess the completeness of the dataset, as missing data can bias results, reduce statistical power, and compromise the validity of conclusions. Table 3.1 summarizes the number of missing values per variable, separately for the European and American datasets. The European dataset exhibits more missing values in absolute terms, but it also contains a greater number of observations.

Variable	Europe	USA
EPS	1058	568
ROA	1835	934
ROE	1191	783
ESG score	1697	781
Environmental score	1697	781
Social score	1697	781
Governance score	1697	781
ESG combined score	1697	781
Market Capitalization	934	447

Table 3.1: Number of Missing Values per Variable for European and American Data (before cleaning).

Interestingly, the ESG-related variables exhibit identical missingness counts within each region. This indicates that ESG information is typically either fully reported or entirely missing, rather than partially missing.

Among the financial ratios, `ROA` has the highest number of missing values in both datasets, implying it may be less consistently reported across firms and time periods. In contrast, `EPS` appears to be more complete, especially in the European dataset, possibly reflecting differences in reporting requirements or data collection processes. `Market Capitalization` also shows a moderate level of missingness, particularly in the European market, which could be due to challenges in obtaining data for smaller or delisted firms. To systematically address missing data in the dataset, the following preprocessing steps were undertaken:

First of all, it was examined whether there are financial years that are containing only missing values for a specific variable, indicating that data for that variable was entirely unavailable during that financial year. Such a pattern could suggest anomalies or reporting issues specific to that period. However, no such years were identified in either the European or American datasets.

As a next step, companies with more than 70% missing values in the ESG variables are removed from the datasets. This results in the exclusion of 41 companies from the European dataset and 11 from the American dataset.

To address excessive missingness in the financial indicators, a similar cutoff was applied to the financial ratios. This leads to the removal of an additional 12 companies from the European dataset and 4 from the American dataset.

The 70% threshold is chosen to balance data quality with keeping enough data for analysis. Removing companies with too much missing data improves the reliability and clarity of the results, while still keeping a large enough sample for meaningful analysis. As noted by Ibrahim et al. (2012) (47), high levels of missing data can significantly affect the precision of estimates and lead to biased conclusions if not handled carefully. Although there's no universal rule for how much missing data is acceptable, research suggests it's important to think carefully about its impact and use practical cutoffs based on the goals of the analysis. Thus, the 70% threshold strikes a reasonable balance, maintaining the reliability of the analysis while minimizing issues caused by missing data.

Subsequently, rows with missing values occurring at the beginning or end of a firm's time series were removed. Specifically, all rows prior to the first non-`NaN` value and all rows following the last non-`NaN` value were excluded. This step ensures that each company's time series starts and ends with actual data, addressing cases where firms may not have reported data during the initial or final years of the period. Since missing data outside of a company's reporting window provide no meaningful information, their removal improves comparability across firms and reduces the influence of incomplete records. The reason missing data outside a company's reporting window appears in the dataset is due to the time range selected when extracting data from LSEG. Not all firms began or ended their ESG and financial reporting in the same years, which results in gaps for companies that were not active or did not report during the full time window.

Additionally, there will be checked if there are companies with random gaps in the data. To assess the reliability of the dataset across companies, there will be investigated whether any company exhibits multiple discontinuities (gaps) in its time series data. Specifically, in this study a gap will be defined as a sequence of more than three consecutive missing values in a given variable, and flagged companies that had more than two such gaps over the full time period. For the European dataset this results in removing five companies, and for the USA dataset this results in removing one company.

Table 3.2 shows the remaining number of missing values after the above procedures. Overall, the number of missing values is relatively low across most variables. Notably, both datasets has a high number of missing values for the `ROA`.

Next, the remaining missing values presented in Table 3.2 were imputed using surrounding valid observations, where possible. Specifically, if a missing value was directly preceded and followed by valid entries, it was replaced with the mean of these two surrounding values. This approach allows for a straightforward imputation that maintains the general trend of the data without introducing strong assumptions.

Variable	**Europe**	**USA**
EPS	10	3
ROA	372	170
ROE	68	69
ESG score	13	12
Environmental score	13	12
Social score	13	12
Governance score	13	12
ESG combined score	13	12
Market Capitalization	16	14

Table 3.2: Number of Missing Values per Variable for European and American Data (after cleaning).

Following this imputation step, the number of missing values decreased significantly. All missing values in the ESG variables were successfully resolved in both datasets. However, some missing values persist in the financial ratios, indicating that imputation using the mean of surrounding values was not always applicable. In the European dataset, missing values remain in all three financial ratio variables: `EPS` still has 3 missing values, `ROA` has 225, `ROE` has 15, and the ESG variables are having 4 missing values. In the American dataset, missing values remain in two financial ratio variables: `ROA` still has 58 missing values, and `ROE` has 12. These remaining gaps will be addressed in the subsequent cleaning steps:

Since the remaining missing values could not be imputed using the surrounding values, this implies that these gaps consist of at least two consecutive missing observations. To better understand the extent of missing data, the maximum number of consecutive missing values was identified for each financial ratio. In the European dataset, the maximum number of consecutive missing values is three for `EPS`, nine for `ROA`, three for `ROE`, and two for ESG variables. For the American dataset, the maximum is zero for `EPS` (as there are no missing values remaining), five for `ROA`, and two for `ROE`. These missing values typically occur in the middle of a company's time series, particularly between the financial years 2010 and 2018.

Because the cause of missingness during this period is unclear and imputing longer gaps would likely introduce unreliable estimates, companies with five or more consecutive missing values were removed. This situation only applied to the variable `ROA`, resulting in the removal of 18 companies from the European dataset and one company from the American dataset.

After applying this filtering step, the maximum number of consecutive missing values was considerably reduced. In the European dataset, no consecutive missing values remain for `EPS`, while the maximum number of consecutive missing values for `ROA` and `ROE` is now four and three, respectively. In the American dataset, the maximum number of consecutive missing values is three for `ROA` and two for `ROE`.

Following the removal of companies with more than five consecutive missing values, the number of missing entries was further reduced. In the European dataset, the variable `ROA` now has 1.77% missing values (108 entities), while `ROE` has only 0.08% missing values (5 entities), and the ESG variables 0.06% (4 entities). In the American dataset, 1.16% (53 entities) of `ROA` values and 0.26% (12 entities) of `ROE` values remain missing.

In order to address the remaining missing values in the financial ratios `ROA` and `ROE`, a custom linear interpolation procedure was implemented. This method was applied individually for each company and consists of the following steps:

First, each time series was scanned for sequences of missing values. Let x_t denote the observed value at year t. A gap was defined as a contiguous block of missing values $[x_{t+1}, \ldots, x_{t+k}]$, bounded by observed values x_t and x_{t+k+1} at either end. That is, both x_t and x_{t+k+1} are known, while $x_{t+1}, \ldots, x_{t+k}$ are missing.

Next, if such a pair of valid boundary values (x_t, x_{t+k+1}) was identified, the missing values were imputed using linear interpolation. The interpolation step size s was computed as:

$$s = \frac{x_{t+k+1} - x_t}{k+1}.$$

Each missing value x_{t+i} for $i = 1, 2, \ldots, k$, the index of a missing value within the gap, was then estimated as:

$$\hat{x}_{t+i} = x_t + i \cdot s.$$

This imputation step will be made clear by the following example:
Suppose for a given company, the `ROA` values over the years 2006 to 2010 are as follows:

Year	ROA
2006	5.0
2007	NA
2008	NA
2009	NA
2010	11.0

Here, we observe a gap of $k = 3$ missing values from 2007 to 2009, with $x_{2006} = 5.0$ and $x_{2010} = 11.0$ as boundary values. The interpolation step size s is computed as:

$$s = \frac{11.0 - 5.0}{3 + 1} = \frac{6.0}{4} = 1.5$$

The missing values are then filled in using the formula $\hat{x}_{t+i} = x_t + i \cdot s$, yielding:

$$\hat{x}_{2007} = 5.0 + 1 \cdot 1.5 = 6.5,$$
$$\hat{x}_{2008} = 5.0 + 2 \cdot 1.5 = 8.0,$$
$$\hat{x}_{2009} = 5.0 + 3 \cdot 1.5 = 9.5.$$

Following this interpolation step, all missing values in both datasets were successfully resolved. This resulted in a final dataset consisting of 6,296 rows for the European dataset and 4,585 rows for the American dataset.

4

Modeling

This chapter presents the modeling approach used in this study. First, the methods for statistical data exploration are described to examine the relationship between ESG scores and financial performance, as well as their associations with control variables, as outlined in Section 4.1. Next, a suitable regression model will be selected and described in Section 4.2, with the aim of predicting ESG performance based on financial metrics, and vice versa. Finally, Section 4.3 presents an optimization model that investigates the trade-off between ESG performance and financial performance.

4.1 Statistical Data Exploration

First, in this section, the data will be tested for normality. Based on the outcome of the normality tests and the variable types, appropriate statistical tests will be selected to examine the relationship between ESG scores and financial ratios, as well as the relationship between the control variables and ESG scores.

4.1.1 Normality Check

One of the most common assumptions in statistical analysis is that the data are normally distributed. Before conducting any statistical tests, it is essential to assess whether this assumption holds. If the data deviate significantly from normality, non-parametric tests, which do not rely on this assumption, may be more appropriate (48).

Therefore, after completing the data cleaning process, a normality check is performed on the ESG variables and financial ratios to determine whether they approximately follow a normal distribution. To assess this, Q-Q (quantile-quantile) plots are examined. A Q-Q plot compares the sample quantiles of the observed data (blue dots) with the theoretical quantiles of a standard normal distribution (red line). The green shaded area represents the 95%-confidence interval. If the points closely follow the red line and lie within the confidence bounds, the data can be considered approximately normally distributed (48).

The Q-Q plots for European companies are shown in Figure 8.3, and for American companies in Figure 8.4. These plots show that the financial performance variables—`ROA`, `ROE`, and `EPS`—exhibit clear violations of the normality assumption.

In contrast, the ESG variables show a more nuanced pattern. The `ESG Score` and `Social Score` aligns reasonably well with the theoretical quantiles in the central range, though minor deviations are visible in the tails, indicating slight skewness. The `Environmental Score` shows moderate deviation, particularly in the lower quantiles, suggesting some positive skewness. The `Governance Score` most closely follows the theoretical quantiles, showing only minor curvature and appearing to be most similar to the normal distribution among the ESG dimensions. The `ESG Combined Score` shows some tail deviations but remains roughly symmetric and can be considered approximately normal. `Market Capitalization`, however, deviates sharply from normality. It exhibits strong right-skewness.

In addition to the visual inspection using Q-Q plots, the normality of the data is also assessed using the Kolmogorov-Smirnov test. This non-parametric statistical method is widely recognized for its power and suitability for large datasets, such as the European dataset, which contains more than 5,000 observations. To ensure consistency across analyses, the Kolmogorov-Smirnov test is applied to both datasets. Specifically, it is used to evaluate the distribution of all ESG variables, financial ratios, and `Market Capitalization`, thereby complementing the graphical analysis.

The Kolmogorov-Smirnov measures the largest discrepancy between the empirical distribution function of the sample and the cumulative distribution function of the reference (49). This discrepancy is quantified by the Kolmogorov-Smirnov statistic (D), which captures the maximum absolute difference between the empirical Cumulative Distribution Function (CDF) and the theoretical CDF, as shown in Equation 4.1.

$$D = \sup_x |F_n(x) - F(x)|, \tag{4.1}$$

where n denotes the sample size (n = 6,296 for the European dataset, and n = 4,585 for the American dataset), $F_n(x)$ is the empirical CDF of the sample, $F(x)$ is the CDF of the reference distribution (50).

Variable	**D-stat (EU)**	**p-value (EU)**	**D-stat (USA)**	**p-value (USA)**
ROA	0.1189	5.23×10^{-78}	0.0985	3.29×10^{-39}
ROE	0.1862	1.46×10^{-191}	0.2402	3.01×10^{-233}
EPS	4071	0.00	0.1585	4.83×10^{-101}
ESG_score	0.0766	1.34×10^{-32}	0.0753	4.57×10^{-23}
Env_score	0.0911	6.44×10^{-46}	0.0989	1.86×10^{-39}
Soc_score	0.0923	3.85×10^{-47}	0.0594	1.60×10^{-14}
Gov_score	0.0591	1.47×10^{-19}	0.0513	6.37×10^{-11}
ESG_Comb_score	0.0477	7.28×10^{-13}	0.0456	1.02×10^{-8}
MarketCap	0.2736	0.00	0.4199	0.00

Table 4.1: Kolmogorov-Smirnov Test Results to check Normality for both datasets (Europe and USA).

For each variable in both datasets, the following null-hypothesis and alternative hypothesis of the Komogorov-Smirnov test are drawn up:

- H_0: The data are normally distributed,
- H_1: The data are not normally distributed.

The test returns a D-statistic and an associated p-value. If the p-value is smaller than the significance level, $\alpha = 0.05$, the null hypothesis of normality is rejected in favor of the alternative hypothesis. This indicates a statistically significant departure from normality, meaning that the data is not normally distributed (49).

The results of the Kolmogorov-Smirnov test for both datasets are summarized in Table 4.1. All variables in the European dataset have p-values below the 5% significance level, leading to the rejection of the null hypothesis of normality for each variable. The same conclusion holds for the USA dataset: all variables reject the null hypothesis at the 5% significance level, indicating that none of the variables are normally distributed. These findings are consistent with the visual inspection of the Q–Q plots, further confirming the non-normality of the data.

4.1.2 Spearman's Rank Correlation test

As shown in the Q-Q plots and confirmed by the Kolmogorov-Smirnov normality tests, none of the continuous numerical variables follow a normal distribution. Given this non-normality, non-parametric methods are more appropriate for further analysis (51).

To examine the relationships between two variables, Spearman's Rank Correlation test can be used. This test is a non-parametric alternative to Pearson's Correlation and is suitable for assessing monotonic relationships between variables without assuming a specific distribution (52).

Spearman's Rank Correlation coefficient r_s is calculated as shown in Equation 4.2.

$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}, \tag{4.2}$$

where d_i is the difference between the ranks of each paired observation, and n is the number of observations.

The hypotheses for the Spearman's Rank Correlation test are as follows:

- H_0: There is no monotonic relationship between the two variables, formally $(r_s = 0)$,
- H_1: A monotonic relationship exists between the two variables, formally $(r_s \neq 0)$.

To determine the statistical significance of the observed correlation, a t-test is conducted using the formula in Equation 4.3:

$$t = \frac{r_s \cdot \sqrt{n-2}}{\sqrt{1 - r_s^2}}, \tag{4.3}$$

where r_s is Spearman's Rank Correlation coefficient and n is the sample size. The resulting t-statistic is used to compute a p-value based on the t-distribution (52). If the p-value is less than the chosen significance level ($\alpha = 0.05$), the null hypothesis is rejected, indicating a significant monotonic association between the two relevant variables.

4.1.3 Kruskal–Wallis Test

In this study, three control variables are considered: `Country of Exchange`, `NAICS National Industry Name`, and `Market Capitalization` (see Section 3.1.1). This section describes the application of the Kruskal–Wallis test to assess whether these control variables are significantly associated with ESG performance.

The Kruskal–Wallis test is a rank-based, non-parametric statistical method used to determine whether there are statistically significant differences in the distribution of a dependent variable (in this case, ESG variables) across three or more independent groups. It is an extension of the Mann–Whitney U test, which is limited to two groups. Unlike parametric tests such as ANOVA, the Kruskal–Wallis test does not assume normally distributed data, making it suitable for this study, since the data violates the assumption of normality (53).

The hypotheses of the Kruskal-Wallis test are defined as follows, where $\tilde{X}_i$ denotes the median of the i-th group:

- H_0: The distributions of the variable of interest are equal across all groups (i.e., the medians are the same across groups). Formally, $\tilde{X}_1 = \tilde{X}_2 = \cdots = \tilde{X}_k$,
- H_1: At least one group differs in its distribution (i.e., at least one group has a different median). Formally, at least one $\tilde{X}_i \neq \tilde{X}_j$ for $i \neq j$,

The expected value of the ranks E_R for each group is given by Equation 4.4.

$$E_R = \frac{n+1}{2}, \tag{4.4}$$

where n is the total number of observations across all groups.

In addition, the variance of the ranks is required to compute the test statistic. The variance of the rank distribution is calculated as in Equation 4.5.

$$\sigma^2 = \frac{n^2 - 1}{12}. \tag{4.5}$$

The Kruskal–Wallis test statistic H is computed as shown in Equation 4.6.

$$H = \frac{n-1}{n} \sum_{i=1}^{k} \left(\frac{n_i \left(\overline{R}_i - E_R\right)^2}{\sigma^2} \right), \tag{4.6}$$

where k is the number of independent groups, n_i is the number of observations in group i, and $\overline{R}_i$ is the average rank of group i.

The Kruskal-Wallis test provides an H-statistic and a corresponding p-value. As with standard hypothesis testing, if the p-value falls below the significance threshold ($\alpha = 0.05$), the null hypothesis is rejected. This implies that the control variable has a statistically significant influence on the variable of interest, for this study the ESG performance.

4.2 Two-Way Predictive Modeling: ESG Scores and Financial Performance

To predict ESG scores and financial performance, a regression model will be developed. Regression analysis is widely used within the *Financial Risk Management* department at KPMG, primarily because it provides interpretable results and clear insights into relationships between variables. Unlike many machine learning methods that often act as "black boxes", regression models offer transparency, an essential feature when working with sensitive financial data for clients. This transparency ensures that predictions and insights can be explained, justified, and aligned with both regulatory requirements and client expectations.

In this section, the analysis will begin with an assessment of multicollinearity, as high correlations among the independent variables can distort regression coefficients and undermine the reliability of the model. Subsequently, the most suitable regression framework will be identified to model the relationship between ESG scores and financial performance, both in predicting ESG scores based on financial indicators and vice versa. The chosen model will then be presented and discussed in detail.

4.2.1 Multicollinearity

Multicollinearity occurs when independent variables in a regression model are highly correlated with each other. This can distort the estimated coefficients, inflate their standard errors, and reduce the model's predictive accuracy. Detecting multicollinearity is therefore essential to ensure reliable and stable regression results.

Two straightforward methods for identifying multicollinearity are scatter plots and the correlation matrix of the independent variables. Scatter plots provide a visual comparison between pairs of variables; a strong linear pattern may indicate multicollinearity. Similarly, the correlation matrix quantifies pairwise relationships, where high correlation coefficients suggest dependencies between variables (54).

Since two regression models are being built, one to predict ESG variables based on financial ratios and the other to predict financial ratios based on ESG variables, both sets of variables serve as independent variables in their respective models. The scatter plots for both datasets, showing all independent variables, are presented in Figures 8.11, 8.12, 8.13, and 8.14.

In the scatter plots of the financial ratios (Figures 8.12 and 8.14), there appears to be no multicollinearity, as evidenced by the wide spread of data points. However, this is not the case for the ESG variables. In Figures 8.11 and 8.13, clear patterns emerge, indicating potential multicollinearity. Specifically, there is a nearly perfect linear relationship between the `ESG Combined Score` and the `ESG Score`, suggesting these variables may be highly correlated.

Furthermore, the `ESG Score` shows somewhat linear relationships with the other ESG variables, indicating possible multicollinearity among these. In contrast, the individual

ESG pillar scores exhibit a wider spread with each other, suggesting that multicollinearity is likely minimal or absent between these pillar scores.

To further investigate multicollinearity, the Variance Inflation Factor (VIF) values are calculated, as defined in Equation 4.7. VIF quantifies how much the variance of an estimated regression coefficient increases due to multicollinearity among the independent variables.

$$VIF = \frac{1}{1 - R_j^2}, \tag{4.7}$$

where R_j^2 represents the coefficient of determination obtained by regressing the j-th independent variable on all other independent variables using linear regression, see Equation 4.8.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}, \tag{4.8}$$

where y_i is the actual value of the dependent variable for observation i, $\hat{y}_i$ is the predicted value of the dependent variable for observation i, $\bar{y}$ is the mean of the observed values of the dependent variable and n is the number of observations (55).

A higher VIF value indicates a greater degree of multicollinearity. While there is no definitive cutoff that classifies a model as 'good' or 'bad' based on VIF, a commonly used rule of thumb is that a VIF of 10 or higher suggests severe multicollinearity (54).

Table 4.2 presents the VIF values for the financial ratios. The consistently low VIF values confirm the visual assessment from the scatter plots, indicating no significant multicollinearity among the financial ratios.

Variable	**VIF_EU**	**VIF_USA**
EPS	1.00	1.04
ROE	1.43	1.25
ROA	1.43	1.24

Table 4.2: VIF of Financial Variables in the EU and USA.

The results for the ESG variables are presented in Table 4.3. Compared to the financial ratios, these VIF values are considerably higher, with the `ESG Score` exhibiting an especially elevated value. This aligns with the correlation matrices of the ESG variables shown in Figures 8.15 and 8.16, where it is evident that `ESG Score` is highly correlated with the other ESG variables. Similarly, the `ESG Combined Score` also shows strong correlations with the other ESG indicators.

Variable	VIF_EU	VIF_USA
ESG_score	62.65	71.24
Env_score	8.61	11.16
Soc_score	17.43	17.97
Gov_score	8.87	9.33
ESG_Comb_score	3.43	3.30

Table 4.3: VIF of ESG Variables in the EU and USA.

Due to the observed multicollinearity, the VIF values were recalculated after excluding the `ESG Score` and `ESG Combined Score`, focusing solely on the ESG pillar scores: `Environmental Score`, `Social Score`, and `Governance Score`. These revised VIF values are presented in Table 4.4. As shown, the VIF values have significantly decreased compared to the previous calculation, indicating little to no multicollinearity among the ESG pillar scores.

Variable	VIF_EU	VIF_USA
Env_score	1.85	2.25
Soc_score	1.96	2.20
Gov_score	1.24	1.26

Table 4.4: VIF of only the ESG pillar scores in the EU and USA.

Therefore, when constructing a regression model with the financial ratios as the dependent variables and the ESG variables as independent variables, only the ESG pillar scores will be used.

4.2.2 Search for an Appropriate Regression Model

This paragraph investigates suitable regression techniques to model the relationship between ESG indicators and financial ratios. Selecting an appropriate regression model is crucial for accurately capturing this relationship. Several modeling approaches are explored based on the characteristics of the dataset.

The analysis begins with a straightforward approach: Ordinary Least Squares (OLS) regression. Scatter plots are generated to evaluate whether linear relationships exist between ESG scores and financial performance (see Figures 8.5 and 8.6). Each plot includes a fitted regression line (in red) along with the corresponding R^2-value. The wide dispersion of data points around the regression lines and the consistently low R^2-values suggest the absence of a meaningful linear relationship. These findings indicate that OLS regression is not suitable for capturing the underlying relationships between ESG scores and financial performance in these datasets.

Subsequently, the distributional properties of the variables are examined. Histograms and skewness values are presented in Figures 8.7, 8.8, and Table 4.5, respectively. The financial ratios exhibit heavy right skewness, particularly in the EU dataset, whereas the ESG variables show mild left skewness.

Variable	**Skewness_EU**	**Skewness_USA**
EPS	11.968	4.399
ROE	5.850	5.802
ROA	1.867	0.880
Gov_score	-0.422	-0.387
ESG_Comb_score	-0.531	-0.142
Soc_score	-0.752	-0.384
ESG_score	-0.771	-0.475
Env_score	-0.848	-0.525

Table 4.5: Skewness of Financial and ESG Variables in the EU and USA.

To mitigate skewness, logarithmic transformations are considered. However, such transformations are valid only if the data approximately follow a log-normal distribution. This assumption was assessed using Q–Q plots (Figures 8.9 and 8.10) and the Kolmogorov–Smirnov test. Both visual inspection and statistical testing rejected the assumption of normality (p-values = 0.00 for all variables), indicating that log transformations and log-linear models are inappropriate.

Given the observed non-linearity and the lack of normality, polynomial regression is considered as an alternative. For this model to be valid, the following assumptions must hold (56):

- **Homoscedasticity**: Residuals should have constant variance and no systematic patterns,
- **Normality of Errors**: Residuals should follow a normal distribution to ensure valid inference,

- **No Multicollinearity**: Predictors should not be highly correlated.

Multicollinearity is addressed in Section 4.2.1 by including only the ESG pillar scores as predictors. Homoscedasticity appears to be satisfied, as evidenced by residual plots in Figures 8.18 and 8.19, where no discernible patterns are present.

However, the assumption of normality of residuals is not met. Q–Q plots (Figures 8.21, 8.22, 8.24 and 8.23) demonstrate significant deviations from the expected normal distribution. This is further confirmed by the Kolmogorov–Smirnov test, which rejects the null hypothesis.

In conclusion, the non-linear, skewed, and non-normally distributed nature of the data underscores the need for a more flexible regression model to capture the complex relationships between ESG scores and financial performance.

4.2.3 Generalized Additive Model

Given the complex, skewed nature of the data and the presence of nonlinear relationships both in the predictors and residuals, a Generalized Additive Model (GAM) is chosen for the regression analysis. GAMs provide a flexible yet interpretable framework that can effectively capture complex relationships in virtually any regression problem (57, 58).

GAM extend the Generalized Linear Model (GLM) framework by incorporating smooth, non-parametric functions of the predictors. While GLMs assume a linear relationship between transformed responses and covariates, GAMs relax this assumption by allowing each predictor to enter the model through its own smooth function. This flexibility enables GAMs to capture complex, nonlinear relationships in the data, while preserving the interpretability of additive models.

A GAM is of the form:

$$\hat{y} = \beta_0 + f_1(x_1) + f_2(x_2) + \cdots + f_m(x_m), \tag{4.9}$$

where $\hat{y}$ is the predicted response, β_0 is the intercept, $f_i(x_i)$ are smooth functions estimated from the data for each predictor x_i, and m the number of predictors.

To model smooth functions in Python, the package `pyGAM` can be used. This package fits a smooth function for each independent variable with respect to the target variable. In `pyGAM`, these smooth functions can take the following forms (59):

- `l()`: linear terms (raw covariates),
- `s()`: spline terms (smooth, nonlinear transformations),

- `f()`: factor terms (categorical variables),
- `te()`: tensor product terms (interactions between variables).

To avoid overfitting, GAMs employ smoothing penalties λ, where $\lambda_j \geq 0$ controls the degree of smoothness. Larger values of λ_j enforce smoother functions by penalizing curvature more heavily, while smaller values permit more flexible fits. However, the flexible fit can result in overfitting.

The smoothing parameter λ plays a critical role in determining the flexibility of the model. In practice, it is often selected through Grid Search by minimizing the Generalized Cross-Validation (GCV) score, which balances model complexity and predictive accuracy. Grid Search is a widely used technique for hyperparameter tuning in machine learning. It systematically evaluates all combinations of specified hyperparameter values to identify the configuration that yields the best model performance. This process performs both model selection and hyperparameter optimization simultaneously.

In Python, the `pyGAM` library offers a convenient method, `gam.gridsearch(X, y)`, which searches over a grid of λ values. The goal is to minimize the GCV score, a performance metric that estimates prediction error while accounting for model complexity. Minimizing the GCV score helps balance bias and variance, thus reducing the risk of overfitting and improving model generalizability (60, 61).

Although GAMs come from a statistical background, their ability to learn the shapes of the smooth functions from data connects them to machine learning. In this sense, GAMs use machine learning techniques to estimate the flexible functional forms of the predictors, often employing penalized splines or other regularized methods to avoid overfitting.

These estimation techniques involve optimization algorithms and data-driven learning, just like in other machine learning models. The difference is that while many machine learning models are often considered "black boxes," GAMs are inherently interpretable. Each smooth function can be visualized, providing insight into how each predictor influences the outcome.

4.3 Optimizing the Trade-Off Between ESG and Financial performance

In this section, the trade-off model and the optimization model will be described. The approach begins with an examination of the trade-offs between the ESG variables and financial performance using trade-off curves. Following this, the concept of Pareto optimality will be introduced to identify combinations of ESG and financial performance that represent optimal trade-offs.

4.3.1 Trade-off

A trade-off occurs when improving one aspect of a system requires compromising another. In other words, increasing one objective often leads to a decrease in another. In spatial conservation planning, it is uncommon for all objectives to be maximized simultaneously, making trade-offs between them a common and necessary consideration (62).

These trade-offs can be quantified and visualized using a trade-off curve, enabling decision-makers to better understand the relationships between competing objectives. This helps in identifying compromise solutions that balance multiple goals effectively (62), see Figure 4.1.

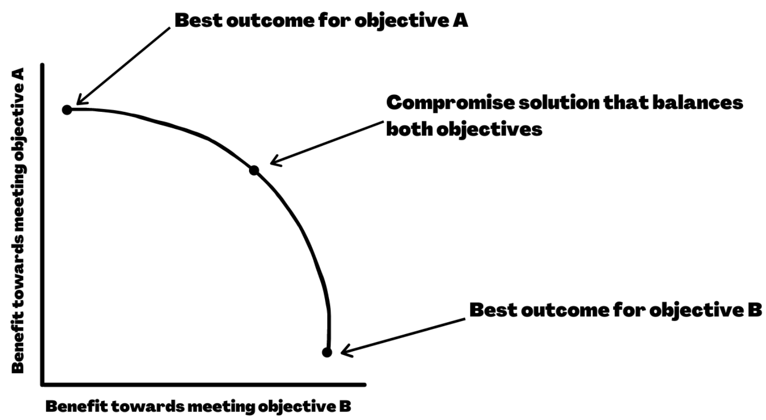


Figure 4.1: Explanation Trade-Off curve (62).

4.3.2 Pareto Optimality Approach

Pareto optimality, named after the Italian economist Vilfredo Pareto (1848–1923), refers to a state in which no individual or objective can be improved without worsening another. It signifies an efficient allocation of resources or outcomes, where all potential mutual gains have been exhausted. While the concept originated in economics, particularly as a cornerstone of welfare economics, it has since been widely adopted in other fields such as engineering, operations research, computer science, and environmental planning. In these contexts, Pareto optimality is used to evaluate trade-offs between competing objectives, such as cost vs. performance or conservation vs. development. Importantly, Pareto optimality does not imply fairness or equity; it simply ensures that no further improvements can be made without incurring a cost elsewhere (63).

Multi-Objective Optimization using NSGA-II

In this study, both ESG performance and financial performance are treated as objectives to be maximized, resulting in a multi-objective optimization problem. To address this, the `Non-dominated Sorting Genetic Algorithm II (NSGA-II)` will be employed. The NSGA-II algorithm, introduced by Deb et al. (2000)(64), is a widely-used evolutionary algorithm for multi-objective optimization. NSGA-II is well-suited for this task due to its ability to efficiently approximate the Pareto front by generating a diverse set of trade-off solutions that reflect the optimal balance between conflicting objectives. It is known for its computational efficiency, elitism mechanism, and its ability to preserve diversity through the use of crowding distance (65).

One of the core components of NSGA-II is fast non-dominated sorting, where the population is divided into different Pareto fronts based on dominance relationships. A solution is considered non-dominated if no other solution is better in all objectives. Another important feature is the crowding distance, which serves as a density estimation metric. It is used to maintain diversity within each front by encouraging a uniform spread of solutions across the Pareto front. Finally, elitism is implemented by retaining the best solutions from both the current and previous generations, ensuring that the overall quality of the Pareto front does not deteriorate over time.

The NSGA-II will be applied using the Generalized Additive Model (GAM) for predicting the ESG performance based on financial performance, and vice versa. The decision variables in this optimization framework include the configuration parameters of the GAM, such as the selection of predictors and the smoothing parameters.

By searching across a wide range of GAM configurations, the NSGA-II algorithm aims to identify a diverse set of models that achieve different balances between ESG compliance and profitability. Because no single configuration can optimize all objectives simultaneously, NSGA-II seeks out the set of Pareto-optimal solutions, those for which no objective can be improved without deteriorating at least one other.

Knee Point

By using multi-objective optimization, several optimal solutions, known as Pareto optimalities, are obtained. To identify the most suitable one, the Knee Point on the Pareto front is determined. In multi-objective optimization, the Knee Point is considered the 'best compromise' solution, as it marks the position where improving one objective would result in a significant deterioration of at least one other. In other words, it represents the optimal trade-off among all objectives.

Formally, the Knee Point is the solution on the Pareto front that yields the highest marginal utility, where any further improvement in one objective leads to a disproportionate loss in another. An illustrative example is shown in Figure 4.2. Intuitively, this point lies in a 'bulge' of the Pareto front, where trade-offs between objectives are most pronounced. When no explicit preference information is available, the Knee Point is often considered the most desirable choice, as it provides a balanced compromise between conflicting goals (66).

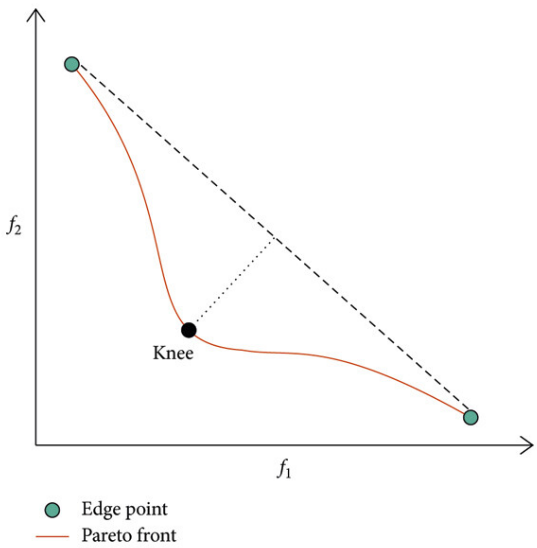


Figure 4.2: Illustration of the Knee Point on the Pareto Front (66).

5

Exploratory Results and Model Development

In this chapter, the results of the exploratory analyses and model development steps, as outlined in Chapter 4, are presented. Section 5.1 begins with a statistical examination of the relationship between ESG scores and financial performance, using Spearman's Rank Correlation test. In Section 5.2, the effects of the control variables—`Market Capitalization`, `NAICS National Industry Name`, and `Country of Exchange`—on ESG performance are assessed through the Kruskal–Wallis test. Next, the predictive capabilities of the Generalized Additive Model (GAM) are evaluated. Section 5.3 focuses on predicting ESG performance based on financial indicators, while Section 5.4 addresses the reverse: predicting financial performance from ESG scores. Finally, Section 5.5 outlines the development and formulation of the optimization model, including the objective function and associated constraints, as a foundation for the final trade-off analysis.

5.1 Relation ESG and Financial Performance

To investigate the relationship between ESG scores and financial ratios, the Spearman Rank Correlation coefficient (described in Section 4.1.2) is calculated between the five ESG variables and three financial ratios. The following hypotheses are tested for each combination of ESG variables and financial ratios:

- H_0: There is no monotonic relationship between the ESG variable and the financial ratio, formally ($r_s = 0$),
- H_1: A monotonic relationship exists between the ESG variable and the financial ratio, formally ($r_s \neq 0$).

The results of the Spearman Rank Correlation test for both datasets are shown in Table 5.1. This table shows that in both the European and USA datasets, the majority of p-values are extremely small (many below 10^{-10}), indicating strong statistical evidence against the null hypothesis. This suggests that monotonic relationships exist between the ESG variables and the financial ratios.

ESG Variable	**Financial Ratio**	**r_s-stat** (EU)	**p-value** (EU)	**r_s-stat** (USA)	**p-value** (USA)
ESG_score	EPS	0.1438	1.96×10^{-30}	0.2513	5.71×10^{-67}
	ROA	-0.1408	2.94×10^{-29}	0.0405	6.15×10^{-3}
	ROE	-0.1162	2.21×10^{-20}	0.0950	1.18×10^{-10}
Env_score	EPS	0.1340	1.27×10^{-26}	0.2432	9.96×10^{-63}
	ROA	-0.2903	1.76×10^{-122}	-0.0341	2.09×10^{-2}
	ROE	-0.2089	4.73×10^{-63}	0.0211	1.53×10^{-1}
Soc_score	EPS	0.1232	1.04×10^{-22}	0.2494	5.46×10^{-66}
	ROA	-0.0410	1.15×10^{-3}	0.0734	6.56×10^{-7}
	ROE	-0.0440	4.75×10^{-4}	0.1171	1.77×10^{-15}
Gov_score	EPS	0.0913	3.86×10^{-13}	0.1425	3.17×10^{-22}
	ROA	-0.1330	3.01×10^{-26}	-0.0360	1.48×10^{-2}
	ROE	-0.0767	1.11×10^{-9}	0.0214	1.48×10^{-1}
ESG_Comb_score	EPS	0.1315	1.14×10^{-25}	0.2165	8.91×10^{-50}
	ROA	-0.0542	1.69×10^{-5}	0.0228	1.23×10^{-1}
	ROE	-0.0490	1.01×10^{-4}	0.0840	1.21×10^{-8}

Table 5.1: Spearman Rank Correlation between ESG variables and financial ratios for Europe and the USA.

5.2 Relation Control Variables and ESG performance

In this section, the Kruskal–Wallis test (as described in Section 4.1.3) is applied to assess whether the control variables — `Market Capitalization`, `NAICS National Industry Name`, and `Country of Exchange` — have a significant influence on ESG performance. Each subsection presents the statistical results corresponding to one of the control variables.

For each control variable, the hypotheses of the Kruskal–Wallis test are defined as follows, where $\tilde{X}_i$ denotes the median ESG score of the i-th group:

- H_0: The distributions of the ESG scores are equal across all groups (i.e., the median ESG scores are the same for all control variable categories). Formally, $\tilde{X}_1 = \tilde{X}_2 = \cdots = \tilde{X}_k$,
- H_1: At least one group differs in the distribution of ESG scores (i.e., at least one group has a different median ESG score). Formally, at least one $\tilde{X}_i \neq \tilde{X}_j$ for $i \neq j$.

5.2.1 Market Capitalization

The control variable `Market Capitalization` is a continuous variable and can not directly be used for the Kruskal-Wallis test. Therefore, the `Market Capitalization` values are divided into bins, as shown in Table 5.2. This process results in the creation of a new column in the dataset, labeled `Market Capitalization Category`, which indicates the bin corresponding to each company's `Market Capitalization`. The `Market Capitalization` values in both datasets are expressed in U.S. dollars.

Category	Market Capitalization Range
Micro Cap	< \$300 million
Small Cap	\$300 million – \$2 billion
Mid Cap	\$2 billion – \$10 billion
Large Cap	\$10 billion – \$200 billion
Mega Cap	> \$200 billion

Table 5.2: Company Size Categories Based on Market Capitalization (67).

The results of the Kruskall-Wallis test to check if ESG scores differ significantly across Market Capitalization categories are shown in Table 5.3. This table reports extremely small p-values for all ESG-related variables, all $p < 0.001$, and therefore below the $\alpha = 0.05$ significance threshold, indicating strong evidence against the null hypothesis of equal medians across firm size categories. These results suggest that differences in ESG scores is significantly associated with company size, proxied by `Market Capitalization`.

5.2.2 NAICS National Industry Name

To evaluate whether ESG performance differs significantly between industries, the Kruskal-Wallis test was applied using `NAICS National Industry Name`. The results, shown in Table 5.4, reveal highly significant differences across industries for all ESG variables in both the European and USA datasets. Specifically, all p-values are extremely small (in

ESG Variable	H-stat (EU)	p-value (EU)	H-stat (USA)	p-value (USA)
ESG_score	1307.98	6.20×10^{-282}	863.54	7.15×10^{-187}
Env_score	1084.52	1.72×10^{-233}	858.27	9.98×10^{-186}
Soc_score	1174.22	6.19×10^{-253}	796.42	2.59×10^{-172}
Gov_score	461.92	1.15×10^{-98}	244.73	9.04×10^{-53}
ESG_Comb_score	553.43	1.86×10^{-118}	510.86	2.11×10^{-110}

Table 5.3: Kruskal–Wallis Test Results for relation ESG Variables and Market Capitalization for European and USA Companies.

most cases effectively 0), indicating a strong rejection of the null hypothesis that the distributions of ESG scores are the same across different industry groups.

ESG Variable	H-stat (EU)	p-value (EU)	H-stat (USA)	p-value (USA)
ESG_score	2273.0807	0.00	1584.9828	7.48×10^{-238}
Env_score	2752.9470	0.00	1634.0144	1.60×10^{-247}
Soc_score	2137.5259	0.00	1583.5521	1.43×10^{-237}
Gov_score	1655.7409	1.49×10^{-230}	1342.9948	1.27×10^{-190}
ESG_Comb_score	1636.7958	6.48×10^{-227}	1227.6492	1.87×10^{-168}

Table 5.4: Kruskal-Wallis Test Results for relation ESG Variables and NAICS Industry for European and USA Companies.

5.2.3 Country of Exchange

The European dataset includes companies from various countries of exchange. This subsection examines whether the `Country of Exchange` influences ESG performance within the European dataset. Since the USA dataset consists exclusively of companies listed in the United States, it is not possible to evaluate the impact of `Country of Exchange` on ESG performance for that dataset; therefore, it is excluded from this analysis.

Table 5.5 presents the results of the Kruskal–Wallis test, which assesses the relationship between ESG performance and `Country of Exchange`. The results indicate that for all ESG-related variables, the p-values are well below the conventional significance level of $\alpha = 0.05$. This provides strong statistical evidence to reject the null hypothesis of equal ESG distributions across countries. These findings suggest that a firm's country of exchange significantly influences its ESG performance.

ESG Variable	H-stat	p-value
ESG_score	482.98	1.30×10^{-92}
Env_score	545.91	6.60×10^{-106}
Soc_score	580.34	3.37×10^{-113}
Gov_score	184.30	1.16×10^{-30}
ESG_Comb_score	436.04	9.91×10^{-83}

Table 5.5: Kruskal-Wallis Test Results for relation ESG Variables and Country of Exchange for European dataset.

5.3 Predicting ESG performance

By applying a Generalized Additive Model (GAM) with ESG variables as the response and financial ratios as predictors, the model takes the following form:

$$\hat{y}_i = \beta_0 + \sum_{j=1}^{3} s_j(X_j), \tag{5.1}$$

where $\hat{y}_i$ denotes the predicted value of ESG variable i, with $i \in \{\texttt{ESG Score}, \texttt{Environmental Score}, \texttt{Social Score}, \texttt{Governance Score}, \texttt{ESG Combined Score}\}$. The variable X_j represents the j^{th} financial predictor (`ROE`, `ROA`, `EPS`), and $s_j(\cdot)$ is a smooth spline function estimated from the data, allowing for non-linear effects.

To determine the optimal degree of smoothness, a Grid Search was conducted to identify the penalty terms (λ). The resulting optimal values for the European and USA datasets are presented in Table 5.6.

ESG Variable	(λ) EU	(λ) USA
ESG Score	0.2512	1.0000
Environmental Score	0.0010	0.6000
Social Score	0.0631	1.0000
Governance Score	0.6000	63.0957
ESG Combined Score	0.2512	1.0000

Table 5.6: Optimal penalty terms (λ) for GAM smoothing per ESG variable.

Using these penalty terms, the GAMs were fitted separately for the EU and USA datasets. Table 5.7 reports the resulting R^2-values.

ESG Variable	R^2 **(EU)**	R^2 **(USA)**
ESG Score	0.0410	0.0963
Environmental Score	0.1059	0.1042
Social Score	0.0242	0.0852
Governance Score	0.0392	0.0398
ESG Combined Score	0.0234	0.0738

Table 5.7: R^2-values for GAM model per ESG variable.

The results indicate that the R^2-values are relatively low across both datasets, all remaining below 0.11. This suggests that the financial ratios explain only a small portion of the variation in the ESG variables. In other words, financial performance indicators alone have limited predictive power for ESG outcomes, implying that other factors play a significant role in shaping ESG scores.

Therefore, a new GAM will be constructed, this time incorporating the control variables: `Country of Exchange`, `NAICS National Industry Name`, and `Market Capitalization Category`. These variables are expected to capture broader contextual and structural influences that may enhance the model's explanatory power for ESG performance, as supported by the statistical analysis in Section 5.2.

To incorporate these categorical variables into the GAM, they were converted into numerical representations using integer encoding, where each unique category was assigned a distinct integer value. For example, each country in the `Country of Exchange` column was mapped to an integer between 0 and 16, reflecting the 17 countries present in the European dataset. A similar approach was applied to the `NAICS National Industry Name` and `Market Capitalization Group` variables, resulting in the encoded variables `Country_code`, `Industry_code`, and `MCap_code`.

This transformation is necessary because the `pyGAM` package, used to fit the GAM, requires all input features to be numerical. The encoded categorical variables are then included in the model using factor terms, enabling the estimation of non-linear and group-specific effects without violating the model's assumptions (68).

While one-hot encoding is a common approach for handling categorical data, integer encoding was deemed appropriate in this context because the `pyGAM` implementation applies separate smooth functions to each encoded category value. As such, the model does not interpret the integer codes as ordinal or continuous variables, but rather as identifiers for distinct categories. This preserves the categorical nature of the variables while maintaining modeling efficiency.

The extended GAM, including the categorical features and interaction terms, for the European dataset is specified as:

$$\begin{aligned}\hat{y}_i = \beta_0 + \sum_{j=1}^{3} s_j(X_j) + \sum_{k=1}^{3} f_k(Z_k) \\ + \sum_{1 \leq a < b \leq 3} te_{a,b}(Z_a, Z_b),\end{aligned} \tag{5.2}$$

where $Z = (\texttt{Industry\_code}, \texttt{MCap\_code}, \texttt{Country\_code})$ are categorical factors modeled by factor smooths $f_k(\cdot)$. The terms $te_{a,b}(\cdot,\cdot)$ represent tensor-product smooth interactions between pairs of categorical variables, capturing their joint effects.

For the USA dataset, since all companies are from a single country, the model reduces to:

$$\begin{aligned}\hat{y}_i = \beta_0 + \sum_{j=1}^{3} s_j(X_j) + \sum_{k=1}^{2} f_k(Z_k) \\ + te_{1,2}(Z_1, Z_2),\end{aligned} \tag{5.3}$$

with $Z = (\texttt{Industry\_code}, \texttt{MCap\_code})$ as categorical factors.

After this, again Grid Search is applied to provide the optimal penalty terms. This resulted in a penalty term for each ESG variable of $\lambda = 0.001$ for the EU and for the USA data $\lambda = 1.0$. One possible reason for the difference in optimal penalty terms is the categorical structure of the datasets. The USA dataset includes only a single country, reducing interaction complexity among categorical variables, whereas the European dataset spans 17 countries, resulting in more granular and complex interactions that require less penalization to be effectively captured. By using these penalty terms and applying the GAM with interaction terms as in Equations 5.2 and 5.3, results in the following R^2-values, see Table 5.8.

ESG Variable	R^2 **(EU)**	R^2 **(USA)**
ESG Score	0.5592	0.5275
Environmental Score	0.6121	0.5168
Social Score	0.5314	0.5119
Governance Score	0.4387	0.3765
ESG Combined Score	0.4417	0.4282

Table 5.8: R^2-values for Extended GAM Models per ESG variable.

This represents a significant improvement compared to the model using only financial ratios. Given that real-world financial data is employed in this study, these R^2-values are notably high.

To further enhance the model, additional interaction terms are incorporated into the GAM. Specifically, the model for the European data is extended to include:

$$\begin{aligned}\hat{y}_i = \beta_0 &+ \sum_{i=1}^{3} s_i(X_i) + \sum_{j=1}^{3} f_j(Z_j) \\ &+ \sum_{1\le a<b\le 3} te_{CC}(Z_a, Z_b) + \sum_{1\le c<d\le 3} te_{NN}(X_c, X_d) \\ &+ \sum_{i=1}^{3}\sum_{j=1}^{3} te_{NC}(X_i, Z_j),\end{aligned} \tag{5.4}$$

where $te_{NN}(\cdot,\cdot)$ are tensor-product interactions between the financial ratios, and $te_{NC}(\cdot,\cdot)$ are tensor-product interactions between financial ratios and categorical control variables.

For the USA data, where only two categorical factors exist, the extended GAM becomes:

$$\begin{aligned}\hat{y}_i = \beta_0 &+ \sum_{i=1}^{3} s_i(X_i) + \sum_{j=1}^{2} f_j(Z_j) \\ &+ te_{CC}(Z_1, Z_2) + \sum_{1\le c<d\le 3} te_{NN}(X_c, X_d) \\ &+ \sum_{i=1}^{3}\sum_{j=1}^{2} te_{NC}(X_i, Z_j).\end{aligned} \tag{5.5}$$

By applying Grid Search for these models, the following smoothing penalty terms are computed, see Table 5.9.

ESG Variable	**λ (EU)**	**λ (USA)**
ESG Score	0.001	0.0158
Environmental Score	0.001	0.0631
Social Score	0.001	0.0040
Governance Score	0.001	0.2512
ESG Combined Score	0.001	0.0631

Table 5.9: Optimal Penalty Terms (λ) for Extended GAM Model per ESG variable.

By applying this GAM with these penalty terms, the following R^2-values were calculated, see Table 5.10. This table shows again an improvement compared to the previous GAM.

ESG Variable	R^2 **(EU)**	R^2 **(USA)**
ESG Score	0.5985	0.5734
Environmental Score	0.6532	0.5555
Social Score	0.5687	0.5633
Governance Score	0.4843	0.4024
ESG Combined Score	0.4882	0.4632

Table 5.10: R^2-values of the Extended GAM model per ESG variable.

However, this improvement is way less than the previous improvement. Since for this study real world financial data is used, these R^2-values are totally fine for a reliable model.

To model ESG scores using a GAM, the response variable ideally needs to be unbounded, i.e., it should range from $-\infty$ to $+\infty$. However, ESG scores are constrained to the interval $[0, 100]$. Directly modeling such bounded data can lead to unrealistic predictions or unstable behavior, especially near the boundaries. In practice, this is indeed observed: the current GAM model produces ESG predictions outside the valid 0–100 range. As a result, the model in its current form is not suitable for use in a client-facing context. To address this, the GAM models defined in Equations 5.4 and 5.5 will be used with the additional constraint that ESG predictions must remain within the $[0, 100]$ range. This is achieved by transforming the ESG scores to the logit scale before modeling and applying the inverse transformation after prediction to ensure outputs are bounded. This results in the following transformation procedure:

- **Rescale to** $(0, 1)$**:** Let $y \in [0, 100]$ denote the raw ESG score. First, rescale and clip the values to avoid numerical instability near the boundaries:
$$\tilde{y} = \text{clip}\left(\frac{y}{100}, \varepsilon, 1 - \varepsilon\right),$$
where ε is a small constant, $\varepsilon = 10^{-3}$, to avoid $log(0)$ and $log(1)$,

- **Logit transformation:** Map the clipped score to the real line using the logit function:
$$z = \text{logit}(\tilde{y}) = \log\left(\frac{\tilde{y}}{1 - \tilde{y}}\right),$$

- **Fit the model:** Train the GAM using z as the transformed response variable,

- **Inverse logit:** Transform the predicted value $\hat{z}$ back to the $(0, 1)$ range using the logistic sigmoid:
$$\hat{\tilde{y}} = \sigma(\hat{z}) = \frac{1}{1 + e^{-\hat{z}}},$$

- **Rescale to** $[0, 100]$**:** Finally, return to the original ESG scale by rescaling:

$$\hat{y} = 100 \cdot \hat{\tilde{y}}.$$

For this purpose, the optimal penalty parameters, obtained via GridSearch, are applied, see Table 5.11. These penalty terms closely resemble those identified in the non-transformed GAM, indicating consistency in the model's regularization behavior.

ESG Variable	λ **(EU)**	λ **(USA)**
ESG Score	0.001	0.0158
Environmental Score	0.001	0.0631
Social Score	0.001	0.001
Governance Score	0.001	0.2512
ESG Combined Score	0.001	0.0631

Table 5.11: Optimal Penalty Terms (λ) for transformed GAM per ESG variable.

Based on this, the corresponding R^2-values are computed, see Table 5.12. These values are highly similar to those obtained from the non-transformed GAM, suggesting that the model's explanatory power remains largely unaffected by the transformation.

ESG Variable	R^2 **(EU)**	R^2 **(USA)**
ESG Score	0.6022	0.5704
Environmental Score	0.5994	0.5204
Social Score	0.5751	0.5666
Governance Score	0.4867	0.4250
ESG Combined Score	0.4894	0.4639

Table 5.12: R^2-values of the transformed GAM per ESG variable.

An alternative approach would be to apply the GAM directly to the raw ESG scores and simply clip the predicted values to the $[0, 100]$ range post hoc. While this method is simpler, it lacks theoretical rigor and may produce biased results near the boundaries. The logit transformation, by contrast, ensures that the bounded nature of ESG scores is respected throughout the modeling process. In this specific case, the difference in predictive accuracy between the two approaches appears limited, suggesting that the transformation did not strongly affect the end results. However, the logit-based approach remains preferable from a modeling perspective, especially in client-facing or high-stakes applications where interpretability and robustness are essential.

5.4 Predicting Financial Ratios

By applying the Generalized Additive Model (GAM) with financial ratios as target variable and the ESG pillar scores as independent variables, the GAM will look like:

$$\hat{y}_i = \beta_0 + \sum_{k=1}^{3} s_k(W_k), \tag{5.6}$$

where $\hat{y}_i$ denotes the predicted value of financial ratio variable i, with $i \in \{\texttt{ROE}, \texttt{ROA}, \texttt{EPS}\}$. The vector $W = (\texttt{Environmental Score}, \texttt{Social Score}, \texttt{Governance Score})$ contains continuous ESG predictors modeled by smooth spline functions $s_k(\cdot)$.

By applying Grid Search, the following optimal penalty terms are calculated, see Table 5.13.

Financial Ratio	**(λ) EU**	**(λ) USA**
ROE	63.0957	251.1886
ROA	63.0957	251.1886
EPS	1000.0000	0.6000

Table 5.13: Optimal penalty terms (λ) for GAM smoothing per Financial Ratio.

By using these optimal penalty terms and applying the GAM, the following R^2-values are occurring, see Table 5.14. Again, really low R^2-values, making this model not reliable.

Financial Ratio	**R^2 (EU)**	**R^2 (USA)**
ROE	0.0298	0.0274
ROA	0.0905	0.0466
EPS	0.0028	0.0274

Table 5.14: R^2-values of the GAM model per Financial Ratio.

Therefore, categorical control variables are also incorporated into the GAM to improve the prediction of financial performance. For the European dataset, the extended GAM specification is given by:

$$\begin{aligned} \hat{y}_i = \beta_0 &+ \sum_{i=1}^{3} s_i(W_i) + \sum_{j=1}^{3} f_j(Z_j) \\ &+ \sum_{1 \le a < b \le 3} te_{CC}(Z_a, Z_b), \end{aligned} \tag{5.7}$$

where $Z = (\texttt{MCap\_code}, \texttt{Industry\_code}, \texttt{Country\_code})$ are categorical predictors modeled by factor terms $f_j(\cdot)$. Interaction terms $te_{CC}(\cdot,\cdot)$ represent tensor-product interactions between pairs of categorical predictors.

For the USA data, where country is not included, the extended GAM becomes:

$$\begin{aligned}\hat{y}_i = \beta_0 &+ \sum_{i=1}^{3} s_i(W_i) + \sum_{j=1}^{2} f_j(Z_j) \\ &+ te_{CC}(Z_1, Z_2),\end{aligned} \tag{5.8}$$

where $Z = (\texttt{MCap\_code}, \texttt{Industry\_code})$ are the categorical predictors for the USA model.

For this extended GAM, the following optimal penalty terms are computed by using Grid Search, see Table 5.15.

Financial Ratio	**(λ) EU**	**(λ) USA**
ROE	3.9811	3.9811
ROA	0.0010	1.0000
EPS	3.9811	1.0000

Table 5.15: Optimal penalty terms (λ) for extended GAM per Financial Ratio.

By applying this GAM, the following R^2-values are occurring, see Table 5.16.

Financial Ratio	R^2 **(EU)**	R^2 **(USA)**
ROE	0.2777	0.2957
ROA	0.6181	0.4905
EPS	0.1818	0.4074

Table 5.16: R^2-values of the Extended GAM per Financial Ratio.

This is again an improvement comparing to the GAM without the control variables.

To further improve the GAM, additional interaction terms are included. The extended model for the European data is specified as:

$$\begin{aligned}\hat{y}_i = \beta_0 &+ \sum_{i=1}^{3} s_i(W_i) + \sum_{j=1}^{3} f_j(Z_j) \\ &+ \sum_{1 \le a < b \le 3} te_{CC}(Z_a, Z_b) + \sum_{1 \le c < d \le 3} te_{NN}(W_c, W_d) \\ &+ \sum_{i=1}^{3} \sum_{j=1}^{3} te_{NC}(W_i, Z_j),\end{aligned} \tag{5.9}$$

where $te_{NN}(\cdot,\cdot)$ is the tensor-product interactions between ESG scores, and $te_{NC}(\cdot,\cdot)$ the tensor-product interactions between ESG scores and categorical control variables.

For the USA model, with only two categorical variables, the equation becomes:

$$\begin{aligned}\hat{y}_i = \beta_0 &+ \sum_{i=1}^{3} s_i(W_i) + \sum_{j=1}^{2} f_j(Z_j)\\ &+ te_{CC}(Z_1, Z_2) + \sum_{1\leq c<d\leq 3} te_{NN}(W_c, W_d)\\ &+ \sum_{i=1}^{3}\sum_{j=1}^{2} te_{NC}(W_i, Z_j).\end{aligned} \tag{5.10}$$

For this the following optimal penalty terms are computed in Table 5.17.

Financial Ratio	**(λ) EU**	**(λ) USA**
ROE	3.9811	3.9811
ROA	0.2512	1.0000
EPS	3.9811	1.0000

Table 5.17: Optimal penalty terms (λ) for extended GAM smoothing per Financial Ratio.

This yields the R^2-values shown in Table 5.18. Although the R^2-values have improved, they remain lower than those obtained when predicting the ESG variables, as shown in Table 5.10, indicating that predicting financial ratios is more challenging and may require additional information.

It is worth noting that the combination of `Country of Exchange` and `Industry` may already explain a substantial portion of the variation in ESG scores. Indeed, these categorical variables capture broad structural and regulatory differences across regions and sectors, which are known to influence ESG performance. Consequently, even relatively simple models that include only these factors could yield meaningful predictions.

However, the extended GAM model presented here aims to go beyond such main effects by capturing potential nonlinear relationships and complex interactions between financial ratios, country, and industry variables. This richer modeling approach is intended to improve explanatory power and better reflect the nuanced determinants of ESG scores observed in the data.

In conclusion, the most extended GAM, which incorporates all relevant interaction terms, achieves the highest R^2-values, establishing it as the most reliable specification in this study. This result aligns with expectations, as including interaction effects enables the

Financial Ratio	R^2 (EU)	R^2 (USA)
ROE	0.3217	0.3477
ROA	0.6630	0.5315
EPS	0.2314	0.4671

Table 5.18: R^2-values of the Extended GAM model per Financial Ratio.

model to capture more complex, nonlinear relationships within the data. Despite this improvement, predicting financial ratios remains challenging, because these outcomes are not solely driven by ESG performance and the control variables. Interestingly, the inverse relationship is stronger: ESG performance is significantly influenced by a firm's financial indicators and contextual factors. This suggests that financial health and firm characteristics play a crucial role in shaping ESG outcomes.

5.5 Optimizing the Trade-Off Between ESG and Financial Performance

This section applies the best-performing predictive models developed in Section 5.3 and 5.4 to optimize the trade-off between ESG performance and financial performance. Specifically, only the Generalized Additive Models (GAMs) that demonstrated sufficient predictive accuracy are used in the optimization process. This ensures that only reliable models are incorporated into the objective functions of the multi-objective optimization framework.

To address multicollinearity among ESG variables, the optimization focuses on the three individual ESG pillars: `Environmental Score`, `Social Score`, and `Governance Score`. These scores are predicted using the GAMs defined in Equation 5.4 for the European dataset and Equation 5.5 for the American dataset.

Regarding financial performance, the GAMs developed for predicting `ROE` and `EPS` showed relatively low R^2-values (see Table 5.18). As a result, these two variables are included in the optimization as independent decision variables rather than as modeled outcomes. This means that both `EPS` and `ROE` are included directly as objective functions to be maximized, based on their raw values rather than GAM-based predictions. In contrast, the GAM for `ROA` demonstrated sufficient explanatory power and is therefore integrated into the optimization model as a predicted objective function.

This results in the following four objective functions, where each function represents the output of a GAM trained to predict the corresponding variable based on inputs $\mathbf{x}$:

$$
\begin{aligned}
f_1(\mathbf{x}) &= \mathrm{GAM}_{\mathrm{Env}}(\mathbf{x}) && \text{(predicted \texttt{Environmental Score})},\\
f_2(\mathbf{x}) &= \mathrm{GAM}_{\mathrm{Soc}}(\mathbf{x}) && \text{(predicted \texttt{Social Score})},\\
f_3(\mathbf{x}) &= \mathrm{GAM}_{\mathrm{Gov}}(\mathbf{x}) && \text{(predicted \texttt{Governance score})},\\
f_4(\mathbf{x}) &= \mathrm{GAM}_{\mathrm{ROA}}(\mathbf{x}) && \text{(predicted \texttt{ROA})}.
\end{aligned}
$$

These objective functions are optimized using the Non-dominated Sorting Genetic Algorithm II (NSGA-II), enabling the exploration of efficient trade-offs between ESG and financial performance.

The objective functions $f_1(\mathbf{x})$, $f_2(\mathbf{x})$, and $f_3(\mathbf{x})$ are defined by the GAMs specified in Equation 5.4 for the European data and Equation 5.5 for the American data. These models use the following input features:

$$\mathbf{x} = [\texttt{EPS},\ \texttt{ROE},\ \texttt{ROA},\ \texttt{country\_code},\ \texttt{industry\_code},\ \texttt{mcap\_code}]\,.$$

The first three components — `EPS`, `ROE`, and `ROA` — are treated as decision variables in the optimization procedure. The remaining features — `country_code`, `industry_code`, and `mcap_code` — are control variables and remain fixed for each run, reflecting real-world constraints such as sector or country classification.

For $f_4(\mathbf{x})$, the GAM defined in Equation 5.9 is used for the European data, and Equation 5.10 for the American data. These models use the ESG pillar scores and control variables as input:

$$
\begin{aligned}
\mathbf{x} = [&\texttt{Environmental\_Score},\ \texttt{Social\_Score},\ \texttt{Governance\_Score},\\
&\texttt{country\_code},\ \texttt{industry\_code},\ \texttt{mcap\_code}]
\end{aligned}
$$

Here, the ESG pillar scores are treated as decision variables in the optimization process, while the control variables are again held fixed. For the American dataset, `country_code` is excluded since all firms are from the same country.

The multi-objective optimization problem aims to simultaneously maximize six objectives: three predicted ESG pillar scores, two financial performance ratios, and the predicted return on assets (`ROA`). Formally, the problem is defined as:

$$\max_{\mathbf{x}\in\mathcal{X}}\ \mathbf{f}(\mathbf{x}) = [f_1(\mathbf{x}),\ f_2(\mathbf{x}),\ f_3(\mathbf{x}),\ \mathrm{EPS},\ f_4(\mathbf{x}),\ \mathrm{ROE}]$$

where $\mathbf{x}$ denotes the decision variables and $\mathcal{X}$ the feasible region.

To ensure that the optimization framework operates within realistic and interpretable boundaries, a set of constraints is imposed on the decision variables for both European and American companies. These constraints keep the optimization grounded in plausible economic and financial conditions, thereby preventing the generation of infeasible or implausible solutions.

$$0 \leq \texttt{Environmental\_Score} \leq 100 \tag{1}$$

$$0 \leq \texttt{Social\_Score} \leq 100 \tag{2}$$

$$0 \leq \texttt{Governance\_Score} \leq 100 \tag{3}$$

$$\begin{cases} 0 \leq \texttt{EPS} \leq 780.29 & \text{(Europe)} \\ 0 \leq \texttt{EPS} \leq 73.6 & \text{(USA)} \end{cases} \tag{4}$$

$$\begin{cases} 0 \leq \texttt{ROE} \leq 529.3 & \text{(Europe)} \\ 0 \leq \texttt{ROE} \leq 612.3 & \text{(USA)} \end{cases} \tag{5}$$

$$\texttt{ROA} \geq 0 \tag{6}$$

Constraints (1)–(3) enforce that each ESG pillar score lies within the ESG reporting range of $[0, 100]$. Although the GAMs are trained to output values within this range, explicit bounds prevent extrapolation when the optimizer proposes input combinations outside the training domain. Constraints (4)–(5) impose dataset-specific bounds on the financial performance variables `EPS` and `ROE`. Both variables are constrained to be non-negative and lie within the historical range observed in the respective datasets. This reflects the focus on financially healthy firms and ensures realistic input to the predictive models. Constraint (6) enforces non-negativity on the predicted return on assets, `ROA`, further supporting the profitability objective.

The variables `industry_code`, `mcap_code`, and `country_code` are not decision variables, but they are included in the GAM input vector and affect the predictions. During optimization, these are held fixed or manually varied to represent specific firm types or segments. No explicit constraints are enforced, but their valid encoding ranges are:

- `industry_code`: 0 to 193 (Europe), 0 to 150 (USA),
- `mcap_code`: 0 to 4 (Europe), 1 to 4 (USA),
- `country_code`: 0 to 16 (Europe only).

These ranges reflect the encoding present in the original datasets and ensure that the model operates within known, interpretable regions.

6

Results: Trade-Offs in Multi-Objective Optimization

This chapter presents the results of the study based on the best-fitted Generalized Additive Model (GAM). First, Section 6.1 analyzes the trade-offs between ESG scores and financial performance. Subsequently, Section 6.2 details the outcomes of the multi-objective optimization process.

6.1 Trade-Off Between ESG and Financial Performance

This section explores the trade-off between ESG and financial performance using the best-performing Generalized Additive Models (GAMs). The trade-off is analyzed by examining the relationships between the dependent (target) variables and the independent (predictor) variables as modeled by the GAMs.

Trade-offs are visualized through trade-off curves, where the x-axis represents one predictor variable and the y-axis represents the corresponding target variable. These curves are directly derived from the fitted GAMs by varying one predictor at a time while holding all others constant at their average values. This approach isolates the marginal effect of each financial ratio on the predicted ESG score (or vice versa). The shaded bands around the curves represent 95%-confidence intervals, indicating the statistical uncertainty of these estimates. These intervals are computed based on the estimated standard errors of the GAM smooth functions: they are calculated as the estimated effect plus or minus 1.96 times the standard error at each value of the predictor, assuming approximate normality. This quantifies the uncertainty due to sample size and model fit.

The shape of the curve reveals the nature of the relationship. A positive slope suggests that increases in the predictor variable are associated with increases in the target variable, implying a synergistic or favorable relationship. A negative or nonlinear slope indicates potential trade-offs or diminishing returns, where improving one variable might adversely affect or have a complex effect on the other. Flat segments indicate ranges where changes in the predictor variable have little or no effect on the target variable.

6.1.1 Results When ESG is the Target Variable

To examine how financial performance influences ESG outcomes for European and American firms, the GAMs specified in Equations 5.4 and 5.5 are utilized. ESG scores are rescaled to a 0–100 range, as detailed in Section 5.3. The resulting trade-off curves for European companies are shown in Figure 6.1. Across all plots, uncertainty tends to increase with higher financial ratios.

A generally positive relationship is observed between `EPS` and all ESG components, suggesting that higher `EPS` values are weakly associated with stronger ESG performance. In contrast, the relationship between `ROA` and ESG variables is nonlinear. For most ESG components, performance initially declines as `ROA` increases, before stabilizing or recovering. The relationship between `ROE` and ESG performance is relatively flat, with minor upward or downward trends depending on the specific ESG dimension. Notably, the `Social Score` exhibits a slight negative trend at higher `ROE` levels, while the other ESG dimensions show a slight increase as `ROE` values rise.

For American firms, the corresponding trade-off curves are presented in Figure 6.2. Again, across all plots, uncertainty tends to increase with higher financial ratios. This figure reveals a strong and consistent positive association between `EPS` and all ESG dimensions. The effect is particularly pronounced for the `Environmental Score` and `Social Score`, both of which increase sharply with higher `EPS`. In contrast, ESG performance tends to decline as `ROA` increases, especially for the `Environmental Score`. The effect of `ROE` on ESG scores is relatively stable, though the `Environmental Score` and `Social score` exhibit mild upward curvature. This suggests that extremely high `ROE` values may be modestly associated with steady or improved ESG outcomes.

Overall, in both datasets, `EPS` shows a positive relationship with ESG performance, indicating that increasing `EPS` is associated with higher ESG scores. `ROA` exhibits an inverse relationship, implying that higher operational efficiency may come at the expense of ESG alignment. Specifically, there is a negative relation between `ROA` and ESG performance

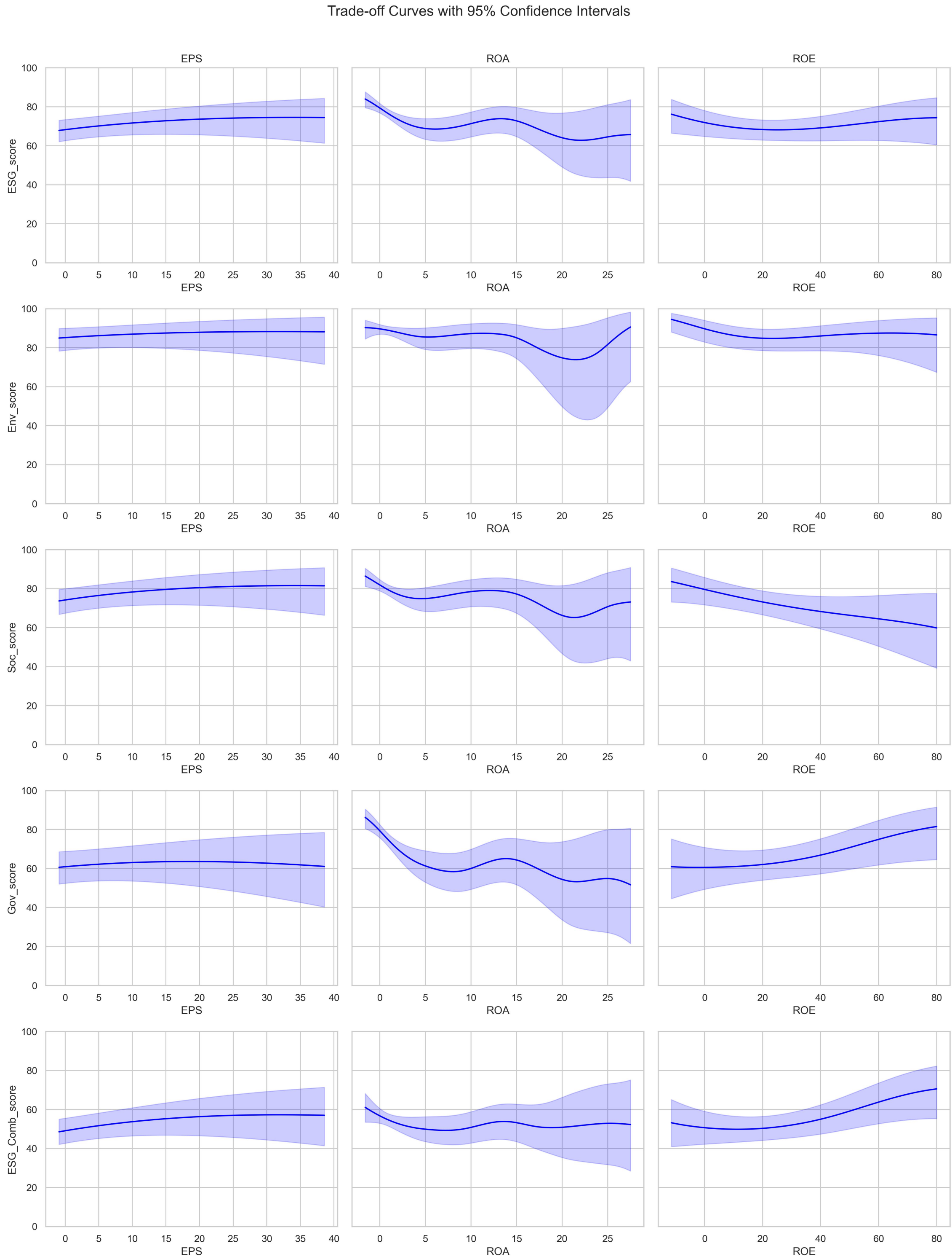


Figure 6.1: Trade-off curves from the GAM with interaction terms for each ESG variable based on Financial Ratios (European data).

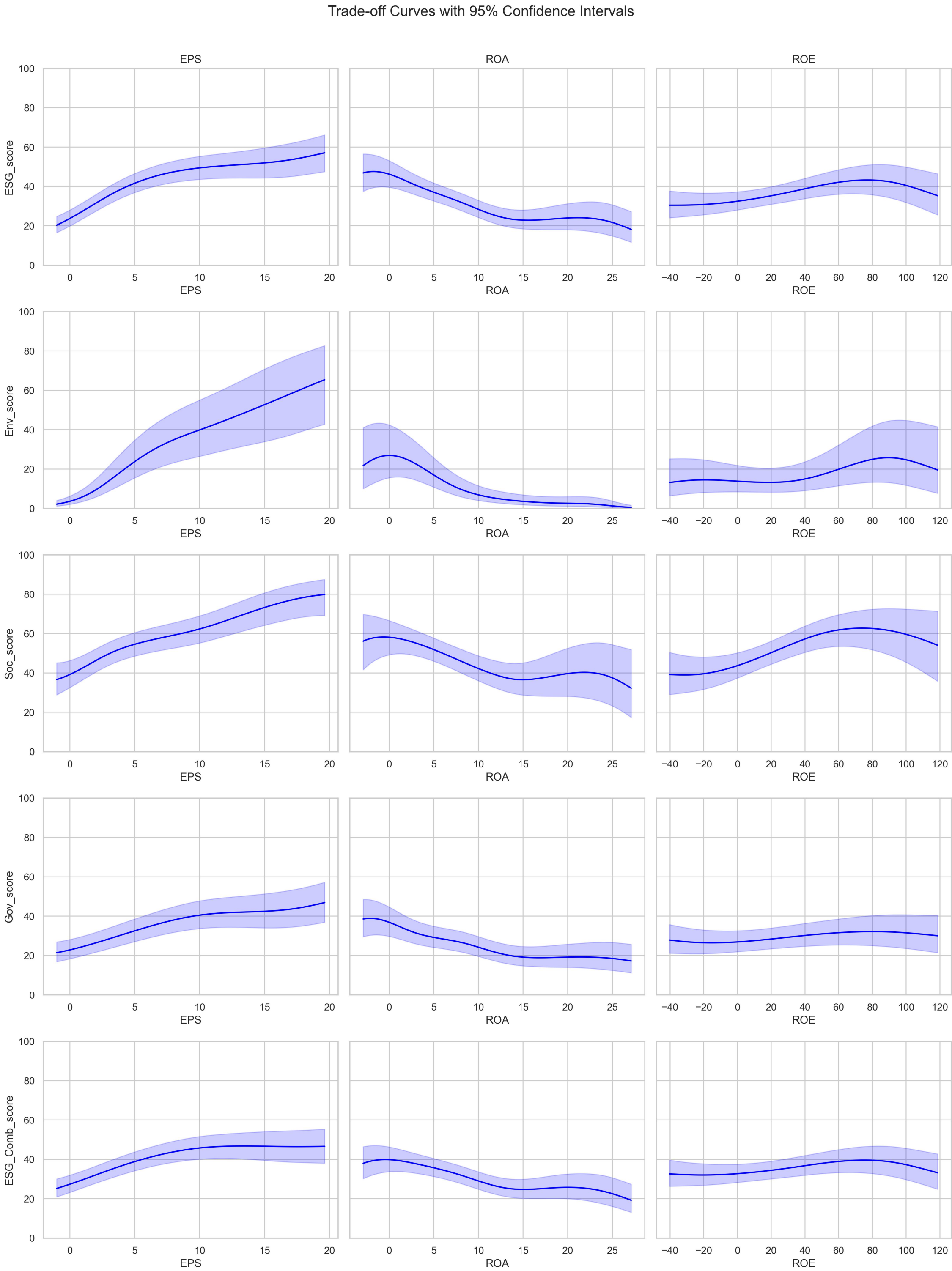


Figure 6.2: Trade-off curves from the GAM with interaction terms for each ESG variable based on Financial Ratios (USA data).

in the USA data, while in the European data, the relationship is more fluctuating and somewhat positive.

6.1.2 Results When Financial Performance is the Target Variable

To investigate how ESG variables influence financial performance for European and American companies, the GAMs described in Equations 5.9 and 5.10 are employed. The resulting trade-off curves for the European dataset are displayed in Figure 6.3. These curves reveal nonlinear patterns and wide 95%-confidence intervals, indicating a high level of uncertainty in the GAM predictions.

The trade-off curves for the European dataset shows that the `Environmental Score` relationship with `EPS` appears nonlinear: `EPS` increases moderately at first but fluctuates across the range, indicating a weak positive association. For `ROA` and `ROE`, a clear downward trend emerges. The `Social Score` also exhibits complex dynamics. `EPS` initially increases but begins to decline after a score of around 50, indicating a nonlinear and ambiguous relationship. `ROA` shows a slight increase followed by a more noticeable decline, while `ROE` fluctuates but generally trends downward near the ESG compliance boundary. Finally, the `Governance Score` shows a more consistently positive relationship with the financial ratios. `EPS` increases steadily with a higher `Governance Score`, indicating a strong positive association. `ROA` remains relatively stable with minor fluctuations, while `ROE` initially increases, stabilizes around a `Governance Score` of 80, and then declines slightly near 90.

Figure 6.4 presents the trade-off curves for the American dataset. Similar to the European results, these curves exhibit nonlinear patterns and wide 95%-confidence intervals, indicating substantial uncertainty in the model estimates.

The trade-off curves for the American data shows that a clear positive relationship emerges between `Environmental Score` and `EPS`. `EPS` steadily increase as `Environmental Score` rise, particularly accelerating beyond a score of 80. In contrast, `ROA` displays a more volatile pattern, relatively stable at moderate `Environmental Score`, but declining at the upper end. `ROE` follows a similar path, remaining stable before trending downward beyond an `Environmental Score` of 70. The `Social Score` shows a generally positive association with `EPS`, increasing fairly linearly across the score range. For `ROA`, however, the relationship is more nuanced: it initially increases but begins to decline after mid-range scores, indicating a potential cost to operational efficiency as social responsibility increases. `ROE` remains relatively stable across the social score spectrum, with only minor fluctuations, suggesting a limited but steady influence on shareholder returns. Lastly, the `Governance Score` reveals mixed results. `EPS` demonstrates a modest increase with

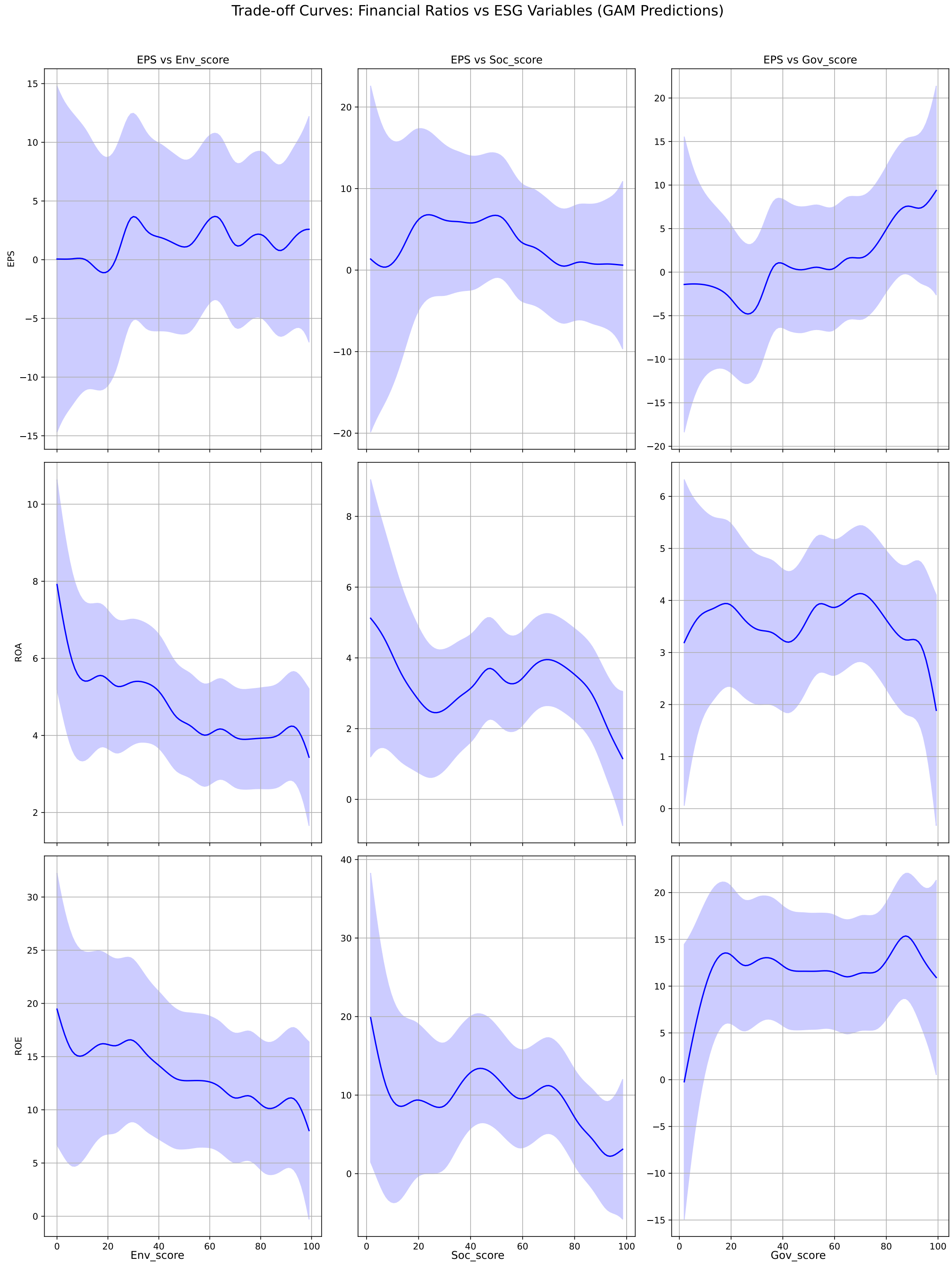


Figure 6.3: Trade-off curves for GAM model with interactions for each Financial Ratio based on the ESG pillar scores for European data.

`Governance Score` up to a point, followed by a decline at very high values. Both `ROA` and `ROE` exhibit downward trends as `Governance Score` increase, with particularly wide variability in `ROE`.

Overall, for both European and American firms, strong `Environmental Score` tends to support higher `EPS`, but may come at the expense of `ROE` and `ROA`. `Social Score` show more ambiguous effects in Europe, while in the USA, they are generally associated with higher `EPS` and more stable profitability. `Governance Score` exhibit the most regional divergence: they are positively related to financial performance in Europe, particularly `EPS`, but in the USA, they appear to offer diminishing returns at higher levels. These findings underscore that ESG strategies may involve trade-offs and that their financial impact is not uniform across geographies or financial ratios.

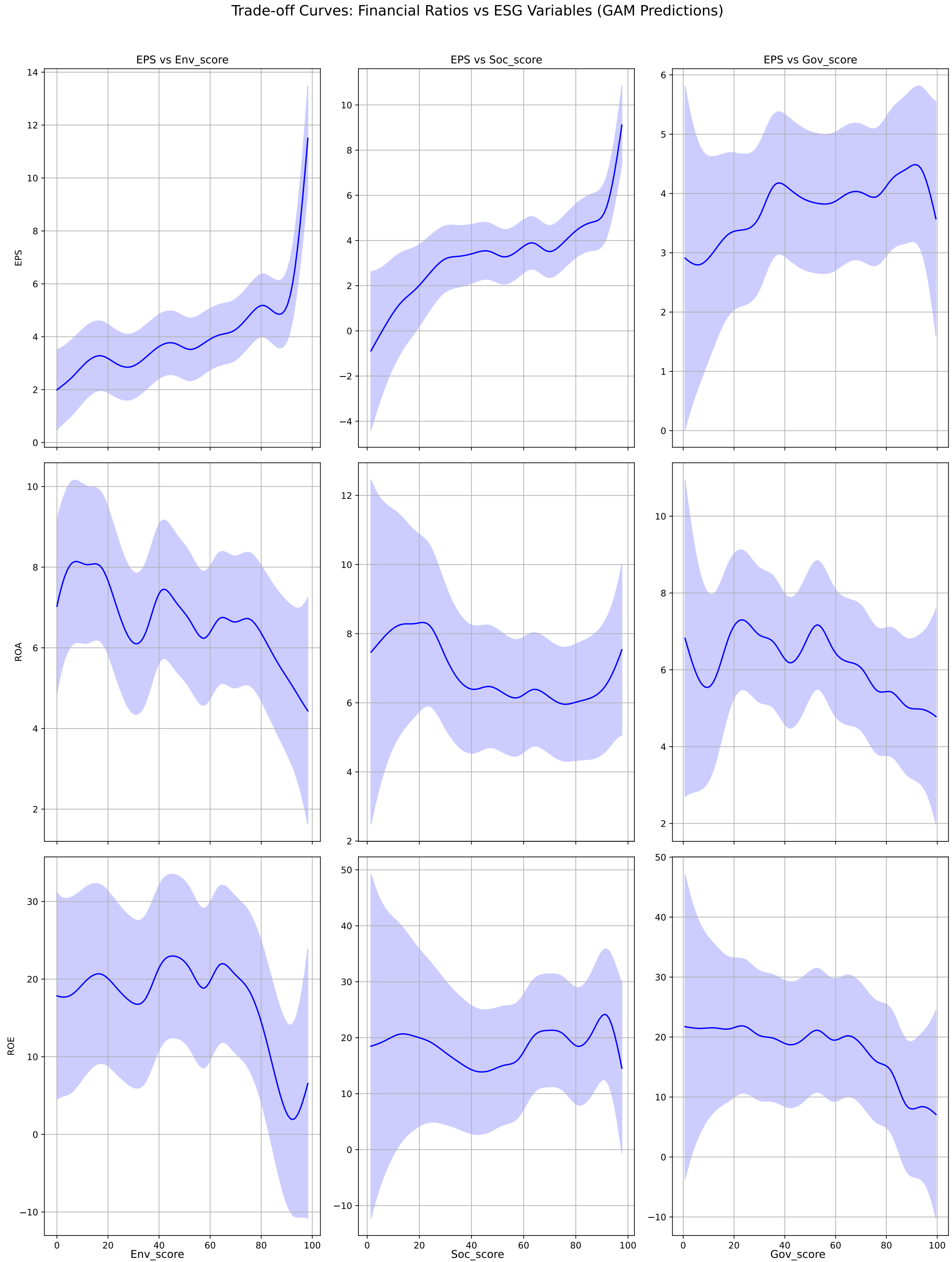


Figure 6.4: Trade-off curves for GAM model with interactions for each Financial Ratio based on the ESG pillar scores for American data.

Summary of the Trade-Offs

To improve the interpretability of the complex relationships between ESG pillar scores and financial performance, the results of the Trade-Off curves are summarized in two separate tables. Table 6.1 presents the trade-offs when ESG pillar scores are modeled as a function of financial ratios, highlighting how financial ratios influence the ESG pillar scores. Conversely, Table 6.2 shows the reverse direction, where financial ratios are predicted using ESG pillar scores, thereby illustrating how ESG compliance may relate to firm-level profitability. The direction and strength of the associations are denoted using symbolic shorthand to indicate positive, negative, flat, or nonlinear relationships. These summaries reveal important patterns and regional differences between European and American firms.

Region	ESG Variable	EPS	ROA	ROE
EU	Environmental	↑ (weak)	↷ / ∼	∼
	Social	↑ (weak)	↷ / ↓	↓ (slight)
	Governance	↑	∼	↑ (slight)
USA	Environmental	↑↑	↓	∼ / ↓ (mild)
	Social	↑↑	↓	∼
	Governance	↑ (mild)	↓	↓ (mild)

Legend: ↑ positive, ↓ negative, ∼ flat/stable, ↷ nonlinear/fluctuating, ↑↑ strong positive.

Table 6.1: Trade-offs: ESG Scores Predicted from Financial Ratios (EU and USA).

Region	Financial Ratio	Environmental	Social	Governance
EU	EPS	↷	↷	↑
	ROA	↓	↓	∼
	ROE	↷ / ↓	↓	↓
USA	EPS	↑↑	↑↑	↑ / ↓
	ROA	↓	↓	↓
	ROE	↓	∼	↓

Legend: ↑ positive, ↓ negative, ∼ flat/stable, ↷ nonlinear/fluctuating, ↑↑ strong positive.

Table 6.2: Trade-offs: Financial Ratios Predicted from ESG Scores (EU and USA).

6.2 Results Pareto Optimization

This section presents the results of the multi-objective optimization using Pareto analysis, as introduced in Section 5.5. The trade-offs between ESG performance and financial metrics are visualized and analyzed. Applying the optimization framework to both the European and American datasets yields a diverse set of candidate solutions, each representing a different trade-off between ESG objectives and financial performance. These solutions are visualized using scatter plots to provide a comprehensive overview of the outcome space. While the visualization offers insight into the trade-off landscape, it should be noted that not all plotted points are strictly Pareto-optimal. Some may be dominated by others, implying that alternative solutions exist that achieve equal or superior performance across all objectives. The red star in the figure highlights the Knee Point, an important solution that balances competing goals and offers a reasonable compromise between maximizing ESG performance and financial performance. It is important to emphasize that the figure does not depict a strict Pareto front, but rather the broader outcome space generated by the optimization process.

Figure 6.5 illustrates the multi-objective solution space for the European dataset. The plots suggest that higher `Environmental Scores` are generally associated with stronger financial performance, particularly in terms of `ROA` and `ROE`. Similarly, the `Social Score` appears positively related to both `EPS` and `ROA`, indicating that firms can achieve solid profitability while maintaining strong social responsibility. The `Social Score` versus `ROE` panel exhibits a more dispersed pattern, though several non-dominated solutions attain both high social performance and high `ROE`, highlighting the potential for mutually reinforcing objectives. In contrast, while the `Governance Score` versus `EPS` panel suggests a generally upward trend, the relationship between `Governance Score` and `ROE` appears inverse. This downward-sloping pattern implies that firms with stronger governance practices may not necessarily exhibit the highest `ROE`, underscoring the complexity of aligning governance performance with short-term profitability metrics.

Figure 6.6 presents the series of scatter plots for American companies. In contrast to the European sample, the figure shows a clearer upward trend in the `Environmental Score` versus `ROE` plot, where higher `Environmental Score` aligns with stronger `ROE`. However, the `Environmental Score` versus `ROA` and `EPS` plots exhibit considerable dispersion, with some optimal points concentrated at lower financial performance levels. This suggests that the link between environmental responsibility and profitability is less consistent in the American context. A distinctive pattern emerges in the `Social Score` versus `ROE` plot,

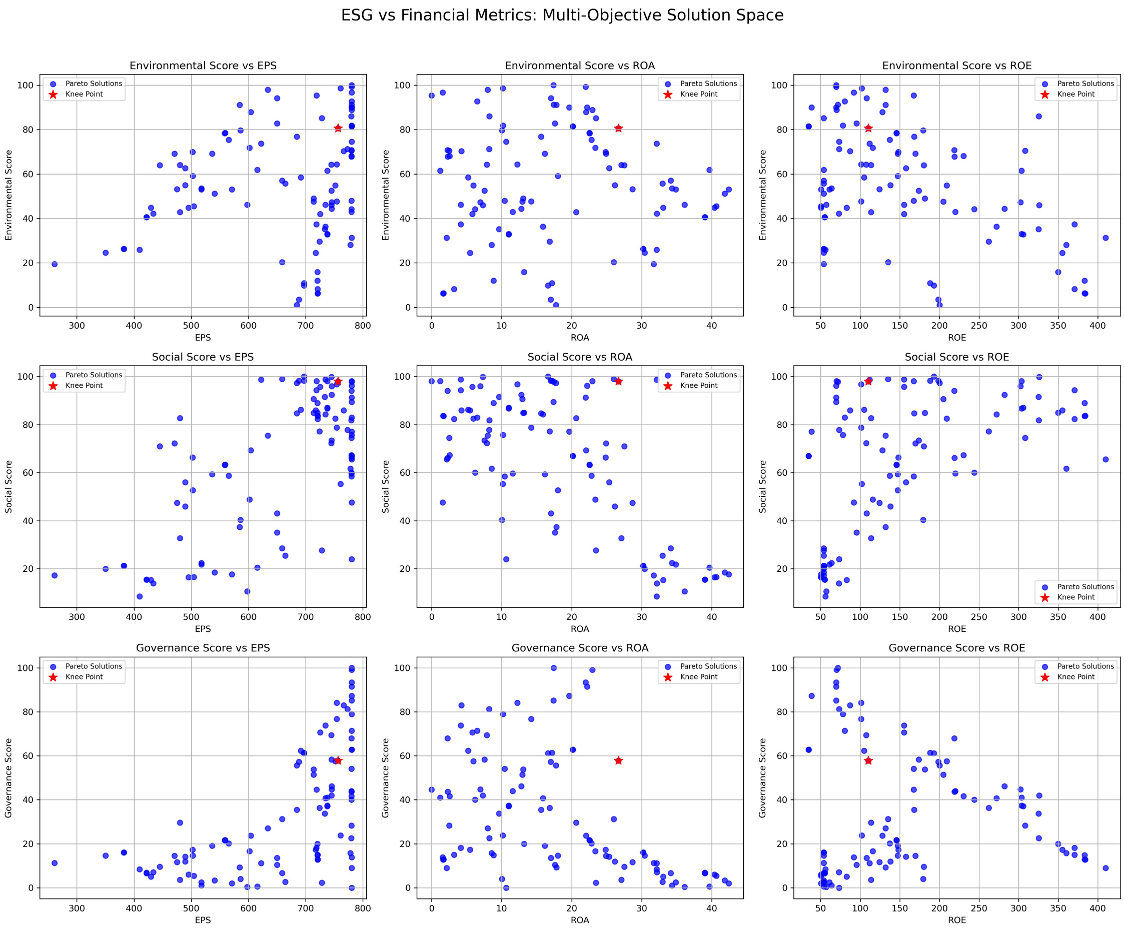


Figure 6.5: Visualization of the multi-objective solution space for the European dataset, illustrating the trade-offs between ESG scores and financial ratios. The red star indicates the Knee Point.

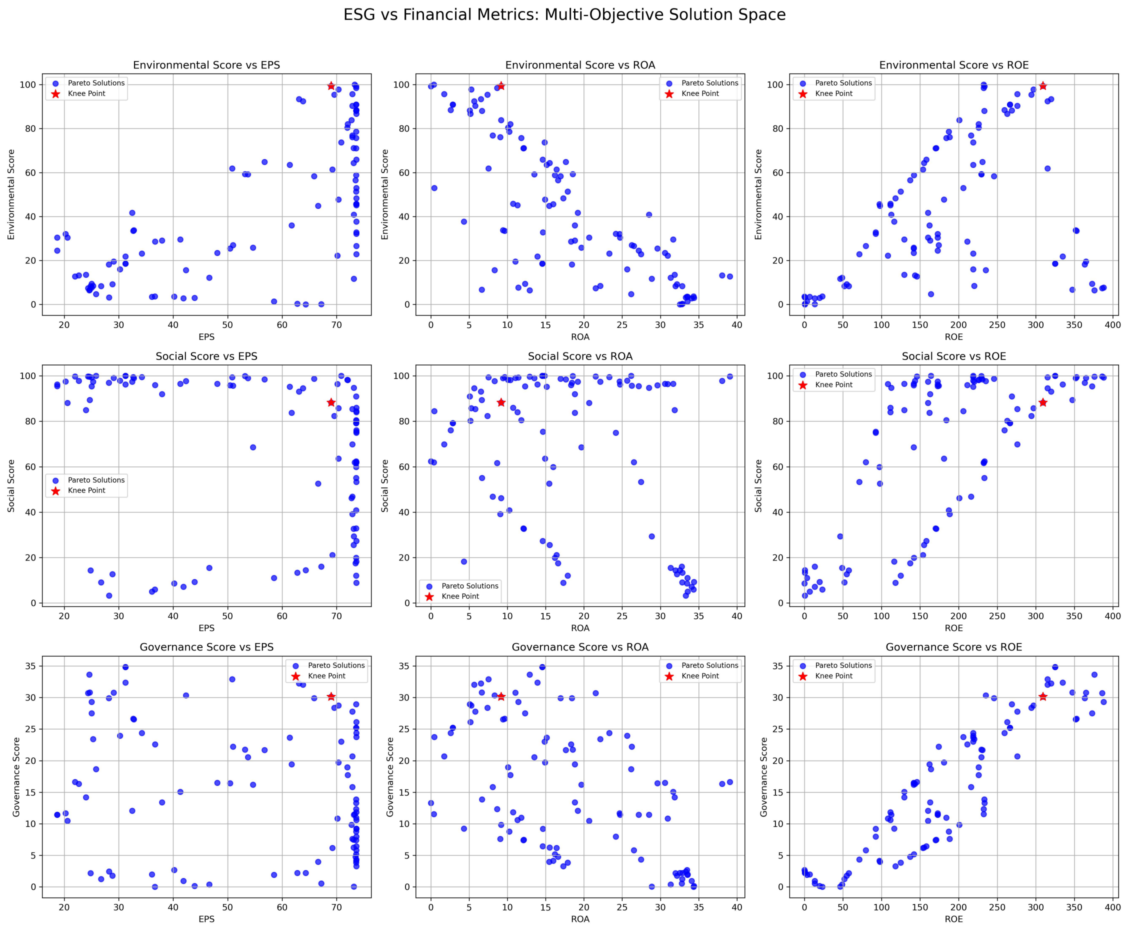


Figure 6.6: Visualization of the multi-objective solution space for the USA dataset, illustrating the trade-offs between ESG scores and financial ratios. The red star indicates the Knee Point.

showing a strong positive relationship between high `Social Score` and `ROE`. The clustering of Pareto-optimal points in the top-right quadrant suggests that high `Social Score` may be a competitive advantage for USA companies. Similar but weaker trends appear in the `Social Score` versus `EPS` and `ROA` panels. The `Governance Score` versus `ROE` panel shows a subtle positive relationship, with some Pareto-optimal solutions combining relatively higher `Governance Score` with strong `ROE`. However, in the `EPS` and `ROA` panels, Pareto-optimal solutions tend to cluster at lower `Governance Score` levels.

Overall, the analysis reveals that relationship between ESG and financial performance vary across regions. In the European dataset, high `Environmental Score` and `Social Score` tend to align with strong financial outcomes, suggesting that ESG integration, especially in these areas, can enhance profitability. In contrast, `Governance Score` presents a more complex picture, where increased transparency and control may temper short-term financial aggressiveness but support long-term stability. In the American data, `Social Score` shows the strongest and most consistent association with financial success, particularly in `ROE`, while the `Environmental` and `Governance` dimensions display weaker or more nuanced relationships. These regional patterns suggest that the strategic value of ESG dimensions is context-dependent: USA markets may prioritize social impact, while European markets reward a more holistic ESG commitment.

The top 10 Pareto-optimal solutions for both the European and American datasets are presented in Table 8.6 and Table 8.7 in the Appendix. These tables report the ESG scores and financial ratios associated with each optimal solution, ranked according to Pareto dominance. Each entry represents a non-dominated solution, meaning no other solution in the dataset simultaneously achieves better or equal performance across all ESG and financial objectives. By analyzing these top-performing trade-off solutions, decision-makers can explore concrete combinations of ESG strengths and financial outcomes.

In addition to the top-performing solutions, the Knee Point (described in Section 4.3.2) is also analyzed. The Knee Point represents the most balanced solution, offering the best compromise between ESG compliance and financial performance. Table 6.3 presents the Knee Point solutions for the European and American datasets.

Variable	**Value (EU)**	**Value (USA)**
Environmental Score	80.59	99.43
Social Score	97.98	88.29
Governance Score	57.84	30.15
EPS	756.51	68.98
ROA	26.68	9.17
ROE	110.19	309.27

Table 6.3: Knee Point Values: ESG Scores and Financial Metrics.

For European firms, the Knee Point yields relatively high `Environmental Score` and `Social Score`, while `Governance Score` is moderate. These ESG outcomes are accompanied by strong financial results, including an exceptionally high `EPS`. This suggests that in the European market, firms can achieve strong ESG compliance without sacrificing, and potentially even enhancing, profitability.

In contrast, the American Knee Point achieves a near-perfect `Environmental Score` and a strong `Social Score`, but the `Governance Score` drops to 30.15. Financially, the USA solution displays lower `EPS` and `ROA` compared to Europe, but a much higher `ROE`. This indicates a different trade-off structure: American firms may deliver strong equity returns (`ROE`) while maintaining high ESG performance in `Environmental` and `Social` dimensions, possibly at the expense of operational efficiency or `Governance` quality.

7

Discussion and Future Research

This chapter presents a discussion of the main findings, including the study's limitations and suggestions for future research. Section 7.1 discusses the key limitations and potential biases that may have influenced the outcomes. Section 7.2 provides recommendations for future research directions to build upon this study.

7.1 Limitations and Biases

Several limitations and potential biases should be considered when interpreting the results of this study. First, the dataset includes only firms based in Europe and the America. As a result, the findings may not be generalizable to companies in other regions. ESG practices and financial market structures vary across countries and cultures, which could influence both ESG adoption and its relationship with financial performance.

Second, the dataset suffers from a sampling bias due to an unequal amount of available data per company. Not all firms report ESG-related metrics consistently over time. Some firms have extensive ESG histories, while others only recently began disclosing such information. This imbalance could skew the results, particularly if firms with more data systematically differ from those with less coverage.

Third, many of the missing values in the dataset are likely informatively missing. In other words, the absence of data is not random but may reflect meaningful structural factors, such as firms not yet existing, not being required to disclose ESG data, or not prioritizing ESG initiatives during certain periods. Ignoring the informative nature of this missingness can lead to biased estimates and misinterpretations.

Another important consideration is the source of ESG scores. ESG ratings are reported by various agencies, each using different methodologies and criteria to assess compliance.

This variation can lead to inconsistencies in how ESG performance is measured across firms. In this study, ESG data is exclusively sourced from the S&P agency, meaning that any bias or limitations inherent to their rating methodology may influence the results. Future studies could consider comparing ESG ratings from multiple providers or using aggregated scores to reduce agency-specific bias.

Additionally, ESG rating methodologies vary significantly across providers. This inconsistency introduces measurement risk and may lead to different conclusions depending on the source of ESG data. A future comparison across multiple data vendors could provide further insight into the robustness of ESG-performance relationships.

Lastly, the limited historical availability of ESG data, especially in the early 2000s, restricts the ability to conduct robust long-term trend analyses. This constraint may limit the study's capacity to capture the evolving role of ESG in financial performance over time.

7.2 Future Research

As global markets continue to evolve, ESG considerations are expected to become an increasingly integral part of corporate governance and investment strategies. Future research could investigate whether firms with higher ESG ratings will demonstrate superior long-term financial performance. It is possible that ESG-related investments may yield financial benefits over a longer horizon, making ESG-focused stocks more attractive to investors.

Furthermore, the ESG scores and financial performance metrics analyzed in this study may be influenced by firm-specific characteristics beyond the scope of the current control variables. While this analysis incorporates `Market Capitalization`, `NAICS National Industry Name`, and `Country of Exchange` as control variables, other factors, such as the level of investment in ESG initiatives, R&D intensity, or management quality, may also play a role in shaping the relationship between ESG compliance and financial performance. Future research could extend the current framework by integrating additional control variables to better isolate the impact of ESG performance and provide a more granular understanding of its drivers.

Additionally, this study does not take time inconsistency into account. That is, the potential delayed effects of ESG compliance on financial performance. In reality, improvements in ESG practices may take several years before reflecting in financial metrics. To address this, future research could incorporate lagged effects into the modeling framework.

The differences between the European and American Knee Points, as presented in Table 6.3, are striking. These contrasts may reflect regional strategic orientations, such as stronger governance structures in European firms and more aggressive profitability focus in the American market. Although further investigation into European variation could offer deeper insight (e.g., comparing northern versus southern European firms).

From a data quality perspective, several limitations in this study could be addressed in future work. First, the dataset contained a large number of missing values, many of which are likely informatively missing. That is, the absence of ESG data may not be due to reporting failure but rather because firms had not yet begun ESG disclosure or did not exist at the time. Future studies should consider modeling these missing values explicitly or employing imputation methods that account for their informative nature.

Moreover, the dataset includes only firms from Europe and the United States. Expanding the geographical scope to include companies from other regions could improve the generalizability and robustness of the findings. Additionally, increasing the number of firms in the dataset would contribute to more reliable statistical inference.

Another area for improvement is outlier detection. In this study, outliers were identified at the company level. However, future work could refine this process by identifying outliers within industry or size groups, as a firm may be an outlier only when compared to its direct peers. Alternatively, robust statistical techniques, such as quantile regression or influence function diagnostics, could be employed to reduce the impact of outliers without the need for arbitrary removal.

There is also limited availability of historical ESG data, especially for the early 2000s. This limits the ability to analyze long-term trends and perform more robust time-series analysis. As ESG reporting continues to improve over time, future studies will be able to draw on richer datasets.

Lastly, methodological enhancements could be made to the modeling approach. For instance, the Generalized Additive Model (GAM) used in this study could be improved by introducing weights to the smooth functions or applying bucketing or discretization before fitting the model.

8

Conclusion

This chapter summarizes the key findings of this thesis and provides an answer to the main research question: *How can the trade-offs between ESG compliance and profitability be quantified in financial decision-making?*

To address this question, the problem was divided into the following sub-questions:

- How can an ESG score be defined, and how can it be used to quantitatively measure ESG compliance and profitability across different firms and industries?
- What is the relationship between ESG scores and financial performance?
- How can a model be developed to analyze and optimize trade-offs between ESG compliance and profitability?

Sections 8.1, 8.2, and 8.3 discuss the findings related to each of these sub-questions. Finally, Section 8.4 synthesizes these findings to answer the main research question and presents the overall conclusion of this thesis.

8.1 Sub-question 1

How can an ESG score be defined, and how can it be used to quantitatively measure ESG compliance and profitability across different firms and industries?

`Environmental`, `Social`, and `Governance` (ESG) scores are composite indicators used to evaluate a firm's performance and commitment across three sustainability dimensions. These scores are typically assigned by specialized rating agencies such as S&P Global, MSCI, Amundi, and Sustainalytics. Each agency applies its own methodology, often combining public disclosures, proprietary data, and expert assessments. A typical ESG assessment includes four key layers:

1. **Availability:** Whether the firm discloses ESG-relevant data.
2. **Quality:** The consistency, accuracy, and reliability of the reported data.
3. **Relevance:** The materiality of ESG data based on the industry context.
4. **Performance:** The extent to which a firm meets ESG benchmarks or standards.

These evaluations are aggregated into dimension-specific scores—`Environmental`, `Social`, and `Governance`—which are further combined into an overall ESG score. For example, S&P uses a 0–100 numerical scale, while MSCI applies a letter-based grading system ranging from `CCC` (lowest) to `AAA` (highest).

To define ESG compliance in quantitative terms, a threshold-based approach is commonly applied. Firms scoring above a certain cut-off, such as 70 on a 0–100 scale, or receiving a rating of `AA` or higher in the MSCI system, are considered to be ESG compliant. These thresholds are typically determined based on best practices in the literature or policy guidelines from regulatory or investment advisory bodies.

In quantitative research, ESG scores can be used as independent variables in statistical and econometric models to study their relationship with financial performance across firms, sectors, and countries. In this study, ESG scores are applied to explain variations in profitability metrics, including Earnings Per Share (`EPS`), Return on Assets (`ROA`), and Return on Equity (`ROE`).

Initial descriptive analysis using Spearman's Rank Correlation suggests a generally positive association between ESG dimensions and the financial ratios. Furthermore, the Kruskal–Wallis tests confirm that the control variables—`Market Capitalization`, `Country of Exchange`, and `Industry`—are significantly associated with ESG performance.

These findings demonstrate that ESG scores are not only useful as summary indicators of sustainability compliance but can also be quantitatively integrated into models assessing firm-level financial performance. By capturing differences in ESG commitment across industries and countries, these scores enable cross-sectional and cross-regional comparisons, providing valuable insights into the trade-offs and synergies between sustainability and profitability.

8.2 Sub-question 2

What is the relationship between ESG scores and financial performance?

The relationship between ESG scores and financial performance is investigated using Spearman Rank Correlation, Generalized Additive Models (GAMs), trade-off curves, and Pareto optimization. The analysis considers bidirectional associations by modeling ESG scores both as dependent and independent variables.

First, the Spearman Rank Correlation coefficients reveal statistically significant monotonic associations between ESG variables and financial ratios. Second, Generalized Additive Models (GAMs) provide insight into the direction and shape of these relationships. When ESG scores are modeled as dependent variables:

- `Environmental Score` shows a consistently positive relationship with `EPS` in both Europe and the USA, suggesting that more profitable firms may invest more in environmental initiatives. `ROA` displays ambiguous patterns, declining with higher environmental scores in the USA, while showing non-monotonic effects in Europe with positive impacts at moderate levels. `ROE` generally exhibits weak or flat associations.
- `Social Score` increases with `EPS` across both regions, again linking profitability to stronger social responsibility. `ROA` tends to decline with higher social scores in the USA, while in Europe the relationship is non-linear, with moderate ESG levels associated with better performance. `ROE` shows weak overall associations, with slight positive trends observed in Europe.
- `Governance Score` is positively associated with `EPS` in both Europe and the USA. European firms show a mild positive correlation between `ROE` and governance, while in the USA, very high `Governance Score` are sometimes linked to lower financial returns. `ROA` shows no consistent trend.

When financial performance is modeled as dependent variables:

- `EPS` remains the most positively and consistently influenced financial metric. Higher `Environmental Score` and `Social Score` in particular are linked with improved `EPS` in both datasets, suggesting a mutually reinforcing relationship where ESG engagement supports profitability.
- `ROA` again reflects a more complex picture. In the USA, higher ESG scores often correspond with lower `ROA`, reinforcing the idea of a performance trade-off. In Europe,

the pattern varies across ESG dimensions, with `Environmental Score` supporting `ROA` at certain levels.

- `ROE` generally exhibits weak responsiveness to ESG inputs, though mild positive effects are observed for European firms with stronger `Governance Score`.

These insights are supported by a Pareto optimization analysis, which evaluates trade-offs between ESG performance and financial outcomes. The resulting Pareto-optimal solutions indicates that firms can achieve high levels of both ESG compliance and financial performance, particularly with respect to `EPS`, where European firms in particular exhibit strong profitability alongside high `Environmental` and `Social Scores`. However, trade-offs emerge for metrics like `ROA` and `ROE`, and the nature of these trade-offs differs between Europe and the USA.

In the European dataset, high `Environmental` and `Social Scores` are positively associated with strong `ROA` and `ROE`, suggesting that ESG integration, especially in these two dimensions, can enhance overall financial performance. Conversely, the relationship between `Governance Score` and profitability appears more complex, with stronger governance sometimes linked to lower `ROE`, potentially reflecting a trade-off between transparency and short-term returns.

In the USA, the trade-off structure differs. Here, a notably strong positive relationship emerges between `Social Score` and `ROE`, indicating that firms with higher social performance may achieve superior equity returns. However, the connection between `Environmental Score` and other financial metrics is less consistent, and `Governance Score` often correlates with lower profitability metrics, particularly `EPS` and `ROA`.

The knee point analysis further highlights these regional differences. European firms at the knee point demonstrate a balanced ESG profile with high `Environmental` and `Social Scores`, moderate `Governance Score`, and outstanding financial outcomes, especially in terms of `EPS` and `ROA`. In contrast, American firms at the knee point achieve near-perfect `Environmental Score` and strong `Social Score`, but with much lower `Governance Score`. These firms show very high `ROE`, suggesting that USA companies may prioritize equity returns over governance structures.

Taken together, these findings underscore that while it is possible for firms to simultaneously achieve high ESG compliance and strong financial performance, the optimal trade-offs vary by region. European firms tend to benefit more uniformly from ESG integration across financial metrics, while American firms may focus more selectively, particularly on maximizing `ROE` through social and environmental initiatives, even if this comes at the expense of governance quality or operational efficiency.

8.3 Sub-question 3

How can a model be developed to analyze and optimize trade-offs between ESG compliance and profitability?

To analyze and optimize trade-offs between ESG compliance and profitability, a multi-objective optimization framework is developed that balances financial performance with ESG outcomes. This framework integrates Generalized Additive Models (GAMs) with the Non-dominated Sorting Genetic Algorithm II (NSGA-II), leveraging the principle of Pareto optimality.

The modeling process begins with estimating ESG compliance as a function of financial performance, and vice versa, using GAMs. This method captures complex, non-linear relationships. To enhance the models, additional contextual variables are incorporated: `Country of Exchange`, `NAICS Industry Name`, and `Market Capitalization Category`. This extended model specification captures the complex, group-specific, and non-linear effects associated with regional, sectoral, and firm-size differences.

Next, the NSGA-II algorithm is employed to perform multi-objective optimization. This evolutionary algorithm identifies optimal solutions for multiple, often conflicting objectives. In this case, the objectives are: Maximize financial profitability and ESG performance. NSGA-II explores a broad design space of GAM specifications, including model structures, variable subsets, and smoothing parameters, to construct a Pareto front: a set of non-dominated solutions where no single objective can be improved without compromising another. Each point on this front represents a different trade-off between ESG compliance and profitability.

To assist decision-makers, the Knee Point of the Pareto front is identified. This point is often considered the most balanced solution, representing the greatest marginal gain across objectives before diminishing returns set in. Selecting this point ensures an efficient trade-off between ESG alignment and financial performance.

8.4 General Conclusion

"How can the trade-offs between ESG compliance and profitability be quantified in financial decision-making?"

This study addresses the research question through a combination of descriptive analysis, statistical modeling, and multi-objective optimization, offering a comprehensive, data-driven answer grounded in empirical evidence.

First, it was established that ESG scores are derived from a combination of firm-reported data and external evaluations by rating agencies. These scores exhibit variation across sectors and regions and can be quantitatively analyzed to assess the extent of sustainability efforts. Descriptive and correlational analyses revealed that ESG scores are positively associated with profitability ratios—`EPS`, `ROA`, and `ROE`—although these relationships are also moderated by control variables such as `Industry`, `Market Capitalization`, and `Country of Exchange`.

Second, through Spearman Rank Correlation analysis and Generalized Additive Models (GAMs), the study uncovered nuanced, often non-linear relationships. For example, `EPS` demonstrated a consistently positive association with ESG scores in both the USA and European datasets. In contrast, `ROA` and `ROE` exhibited more ambiguous or region-specific patterns, with trade-offs particularly evident in the USA market. These insights were further enriched through Pareto optimization, which confirmed that while some firms can achieve both high ESG compliance and strong profitability, such dual performance is not uniformly attainable across all financial metrics.

Finally, a hybrid modeling framework was developed by integrating interpretable GAMs with the NSGA-II algorithm for multi-objective optimization. This approach enabled the construction of Pareto-efficient frontiers that visualize the trade-offs between ESG performance and financial returns. The identification of Knee Points along these frontiers offered practical insights for decision-makers aiming to balance sustainability goals with financial performance.

In conclusion, this research demonstrates that the trade-offs between ESG compliance and profitability can be effectively quantified through a structured and data-driven methodology. ESG scores are not merely symbolic indicators; they carry measurable implications for financial outcomes. However, the strength and direction of these relationships vary depending on the specific financial indicator, regional context, and firm characteristics.

References

[1] M. Friedman. **A Friedman doctrine– The Social Responsibility of Business Is to Increase Its Profits**. *The New York Times*, 1970. 1, 10

[2] R. E. Freeman. **Strategic management: A stakeholder approach**. *Cambridge University Press. The Edinburgh Building*, 1984. 2, 13

[3] S. M. How, C. G. Lee, and D. M. Brown. **Shareholder Theory Versus Stakeholder Theory in Explaining Financial Soundness**. *International Advances in Economic Research*, page 133 – 135, 2019. 2

[4] H. Xinyi. **ESG and Financial Constraints**. *SHS Web of Conferences*, 2023. 2, 3

[5] T. Whelan, U. Atz, T. Van Holt, and C. Clark. **ESG and Financial Performance: Uncovering the Relationship by Aggregating Evidence from 1,000 Plus Studies Published between 2015 – 2020**. *NYU Stern Center for Sustainable Business*, 2021. 2, 11

[6] A. Di Giuli and L. Kostovetsky. **Are red or blue companies more likely to go green? Politics and corporate social responsibilit**. *Journal of Financial Economics*, 2014. 2, 11

[7] T. T. Li, K. Wang, T. Sueyoshi, and D. D. Wang. **ESG: Research progress and future prospects**. *Sustainability*, 2021. 7

[8] **Understanding the "E" in ESG**. *S&P Global*, 2019. 7

[9] R. Carson. *Silent Spring*. Houghton Mifflin, 1962. 7

[10] J. Neilan, P. Reilly, and G. Fitzpatrick. **Time to Rethink the S in ESG**. *Harvard Law School forum on Corporate Governance*, 2020. 8

[11] **What is the "G" in ESG?** *S&P Global*, 2020. 8

[12] S&P Global Sustainable 1. *S&P Global ESG Scores Methodology.* 2024. 8, 9

[13] *ESG Ratings Methodology.* MSCI ESG RESEARCH LLC, 2024. 8, 9

[14] *Amundi Sustainable Finance Disclosure Statement.* Amundi, 2023. 8, 9

[15] *THE ESG RISK RATING: FREQUENTLY ASKED QUESTIONS – FOR COMPANIES.* Sustainalytics, 2020. 8, 9

[16] D. C. Esty and T. Cort. **Values at work: Sustainable investing and ESG reporting**. *Springer*, 2020. 9

[17] A. A. Nabi. **Earnings per share impact on non-financial firms Peroformance**. *Journal of Economic Info*, 2014. 10

[18] D. Cooper, M. Franklin, and P. Graybeal. **Principles of Accounting**. *OpenStax, Rice University*, 2023. 10

[19] B. Chakkravarthy, F. G. Irudayasamy, R. Elangovan, N. Rengaraju, and S. Parayitam. **Relationship between return on assets and firm value: Institutional holdings and firm size as moderators**. *Quality & Quantity*, 2023. 10

[20] D. Purnamasari. **The effect of changes in return on assets, return on equity, and economic value added to the stock price changes and its impact on earnings per share**. *Research journal of finance and accounting*, 2015. 10

[21] A. S. Ramadhan and A. Nuraliati. **THE EFFECT OF NET PROFIT MARGIN AND EARNING PER SHARE ON SHARE PRICES**. *Jurnal Manajemen dan Bisnis*, 2020. 10

[22] D. D. Lee, R. W. Faff, and K. Langfield-Smith. **Vexing Question: Does Superior Corporate Social Performance Lead to Improved Financial Performance?** *Australian Journal of Management*, 2009. 11

[23] J. Nollet, G. Filis, and E. Mitrokostas. **Corporate Social Responsibility and Financial Performance: A non-linear and disaggregated approach**. *Economic Modelling*, 2016. 11

[24] S. Brammer, C. Brooks, and S. Pavelin. **Corporate Social Performance and Stock Returns: UK Evidence from Disaggregate Measures**. *Financial Management*, 2006. 11

[25] M. L. Barnett and R. M. Salomon. **Beyond dichotomy: the curvilinear relationship between social responsibility and financial performance**. *Strategic Management Journal*, 2006. 11

[26] G. Friede, T. Busch, and A. Bassen. **ESG and financial performance: aggregated evidence from more than 2000 empirical studies**. *Journal of sustainable finance & investment*, 2015. 11

[27] C. Zhao, Y. Guo, J. Yuan, M. Wu, D. Li, Y. Zhou, and J. Kang. **ESG and corporate financial performance: Empirical evidence from China's listed power generation companies**. *Sustainability*, 2018. 12

[28] R. Almeyda and A. Darmansya. **The influence of environmental, social, and governance (ESG) disclosure on firm financial performance**. *IPTEK Journal of Proceedings Series*, 2019. 12

[29] S. Y. Sharon, H. Jr-Wei, C. Hong-Yi, and T. Min-Hung. **Detecting corporate ESG performance: The role of ESG materiality in corporate financial performance and risks**. *North American Journal of Economics and Finance*, 2025. 12

[30] O. Ahmed and B. A. Khalaf. **The impact of ESG on firm value: The moderating role of cash holdings**. *Elsevier*, 2025. 12

[31] M. Statman and D. Glushkov. **The Wages of Social Responsibility**. *Financial Analysts Journal*, 2009. 13

[32] M. L. BARNETT. **STAKEHOLDER INFLUENCE CAPACITY AND THE VARIABILITY OF FINANCIAL RETURNS TO CORPORATE SOCIAL RESPONSIBILITY**. *The Academy of Management Review*, 2007. 13

[33] M. Aydoğmuş, G. Gülay, and K. Ergun. **Impact of ESG performance on firm value and profitability**. *Borsa Istanbul Review*, **22(S2)**:S19 – S25, 2022. 13

[34] Teoh T. T., Heng Q. K., J. J. Chia, Shie J. M., Liaw S. W., Yang M., and Nguwi Y. Y. **Machine Learning-based Corporate Social Responsibility Prediction**. *IEEE Xplore*, 2019. 13

[35] S. Kotsantonis and G. Serafeim. **Four Things No One Will Tell You About ESG Data**. *Journal of Applied Corporate Finance*, **31(2)**:50–58, 2019. 14

[36] **Environmental, Social and Governance Scores**. *https://www.lseg.com/en/data-analytics*, 2025. 15, 17

[37] **Fact Sheet: OSSIAM STOXX EUROPE 600 ESG EQUAL WEIGHT NR UCITS ETF 1C (EUR)**. *Ossiam*, 2025. 15

[38] **Fact Sheet: SPDR S&P 500 ESG ETF**. *EFIV*, 2024. 16

[39] **ESG score predictor: A quantitative approach for expanding company coverage**. *Moody's*, 2021. 17

[40] **LSEG ESG Scores**. *LSEG DATA & ANALYTICS*, 2023. 22

[41] F. Berg, J. F. Julian F Kölbel, and R. Rigobon. **Aggregate Confusion: The Divergence of ESG Ratings**. *Review of Finance*, 2022. 22

[42] What are outliers in the data?, Engineering, ENGINEERING STATISTICS HANDBOOK, National Institute of Standards and Technology. 22, 23

[43] D. M. Hawkins. **Identification of outliers**. *Chapman and Hall*, 1980. 22

[44] V. Agarwal. **Outlier detection with Boxplots**. *Medium*, 2019. 22

[45] C. C. Aggarwal. *Outlier Analysis*. Springer, 2017. 23

[46] A. Horsch. **Detecting and Treating Outliers In Python – Part 3**. *Data Science*, 2021. 23

[47] J. G. Ibrahim, H. Chu, and M. H. Chen. **Missing Data in Clinical Studies: Issues and Methods**. *Journal of Clinical Oncology*, 2012. 25

[48] **Normality test**. *Graz, Austria*, 2025. 29, 30

[49] V. Tresvisan. **Comparing sample distributions with the Kolmogorov-Smirnov (KS) test**. *Medium*, 2022. 30, 31

[50] S. Lee. **Practical Insights: Applying Kolmogorov-Smirnov Test in Data Analysis**. *NumberAnalytics*, 2025. 30

[51] **Parametric and non-parametric tests**. *Graz, Austria*, 2025. 31

[52] **Spearman's rank correlation coefficient**. *Graz, Austria*, 2025. 32

[53] **Kruskal-Wallis-Test**. *Graz, Austria*, 2025. 32

[54] J. Murel and E. Kavlakoglu. **What is multicollinearity?** *IBM*, 2023. 34, 35

[55] D. Erika. **Looking at R-Squared**. *Medium*, 2019. 35

[56] A. Sengupta. **Polynomial Regression: A Comprehensive Guide**. *Medium*, 2024. 37

[57] M. Gupta. **Understanding Generalized Additive Models (GAMs) for regression with mathematics**. *Medium*, 2022. 38

[58] I. Shen. **Building interpretable models with Generalized additive models in Python**. *Medium*, 2018. 38

[59] A Tour of pyGAM, pyGAM, 2018. 38

[60] M. Clark. **Generalized Additive Models**. *GitHub*, 2022. 39

[61] **What is Grid Search?** *Dremio*, 2025. 39

[62] **Understanding trade-offs and trade-off curves.** *MARXAN*, 2022. 40

[63] E. Rasure and M. Logan. **Pareto Efficiency Examples and Production Possibility Frontier**. *Investopedia*, 2025. 41

[64] K. Deb, S. Agrawal, A. Pratap, and T. Meyarivan. *A Fast Elitist Non-dominated Sorting Genetic Algorithm for Multi-objective Optimization: NSGA-II*. Springer, Berlin, Heidelberg, 2000. 41

[65] **Non-Dominated Sorting Genetic Algorithm 2 (NSGA-II)**. *GeeksforGeeks*, 2024. 41

[66] Tao Zhang Shengjun Huang Wenhua Li, Guo Zhang. **Knee Point-Guided Multiobjective Optimization Algorithm for Microgrid Dynamic Energy Management**. *Complexity*, 2020. 42

[67] J. Fernando. **Market Capitalization: What It Means for Investors**. *Investopedia*, 2024. 45

[68] E. Pedersen, Z. Taranu, and C. F. Dussault. **GAM with interaction terms**. *Québec Centre for Biodiversity Science*, 2023. 48

Appendix

Country	**Number of Companies**
United Kingdom	105
France	69
Germany	52
Sweden	45
Switzerland	42
Italy	33
Netherlands	24
Spain	22
Denmark	21
Finland	18
Belgium	13
Norway	12
Austria	8
Ireland	6
Portugal	5
USA	2
Poland	2

Table 8.1: Number of companies per country in the European dataset.

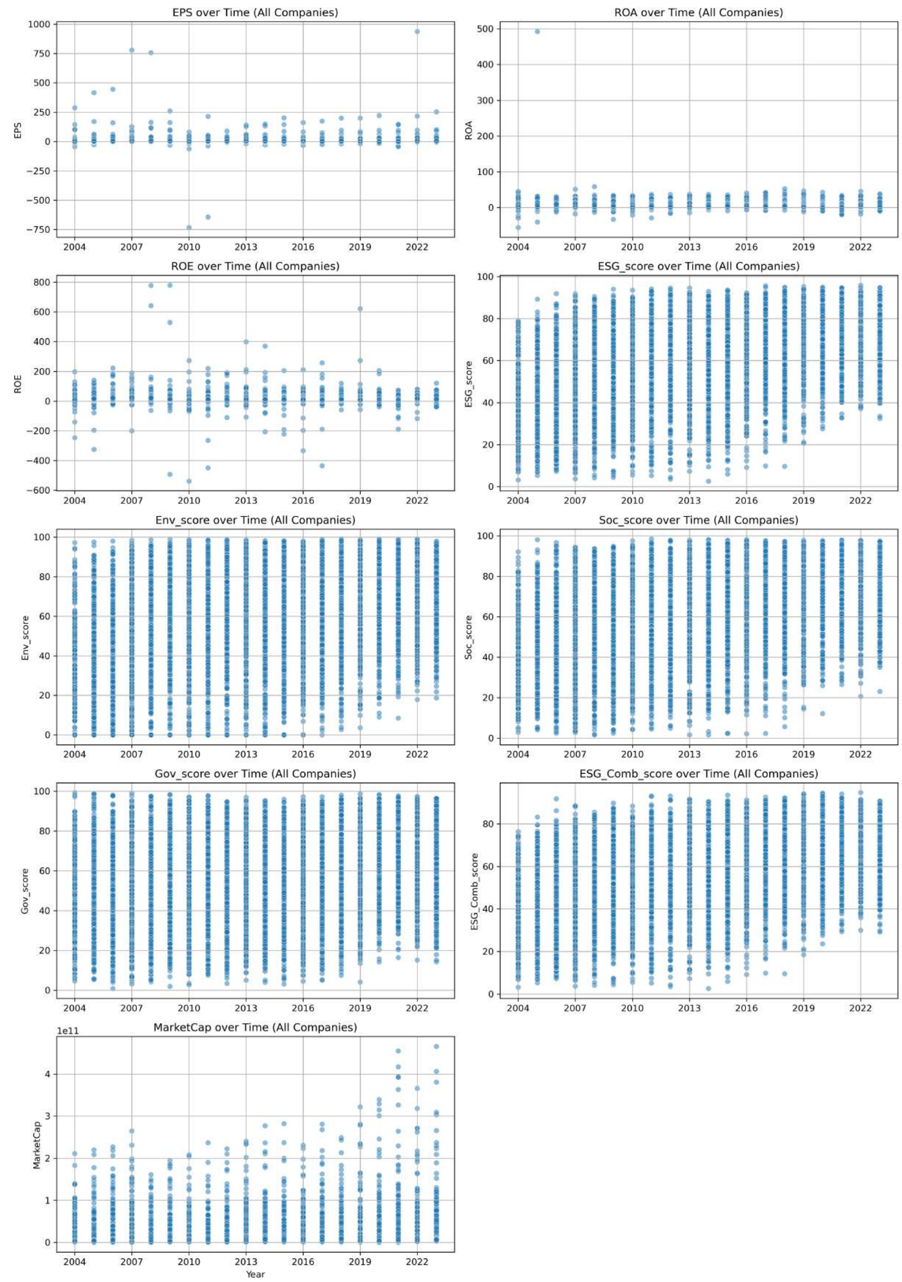


Figure 8.1: Scatter Plot of all variables over time for European dataset.

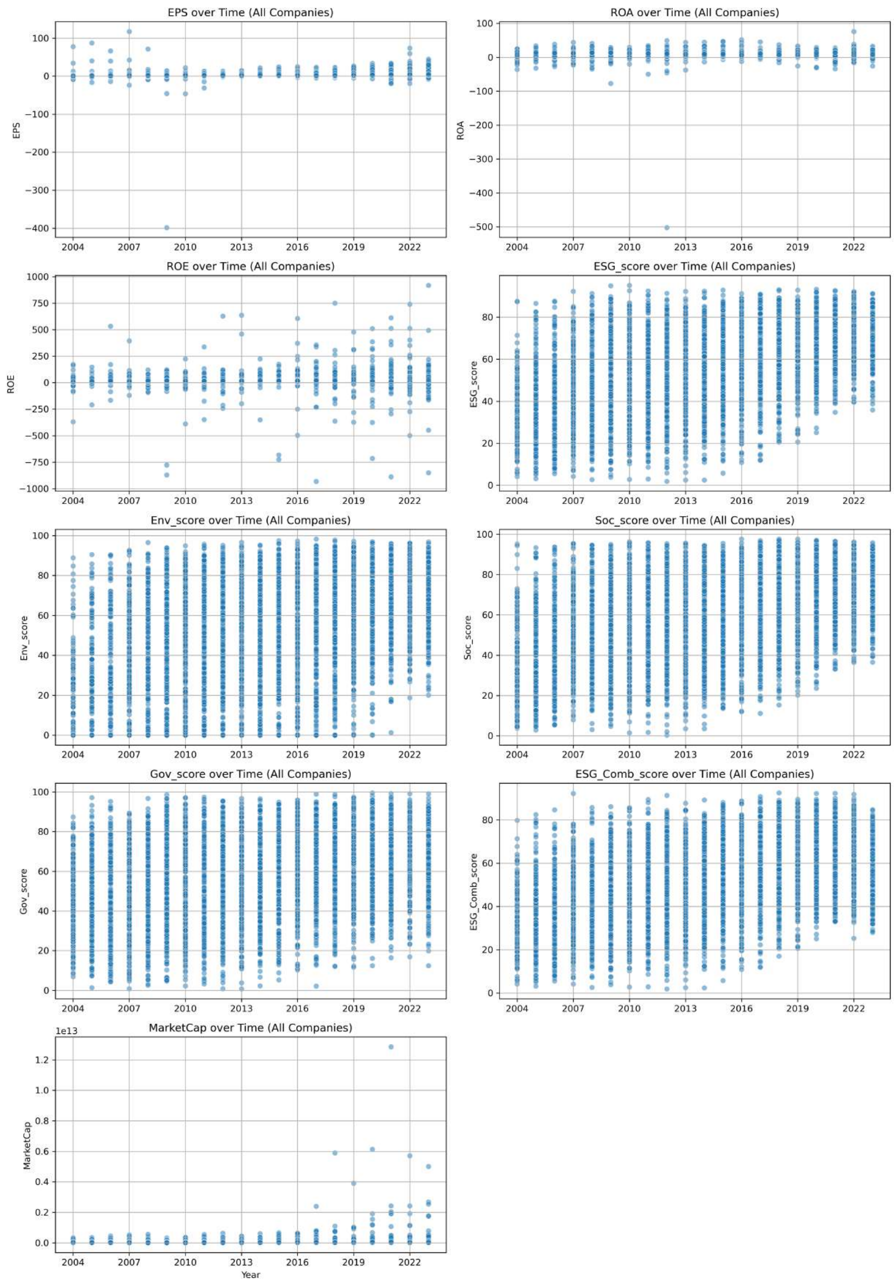


Figure 8.2: Scatter Plot of all variables over time for American dataset.

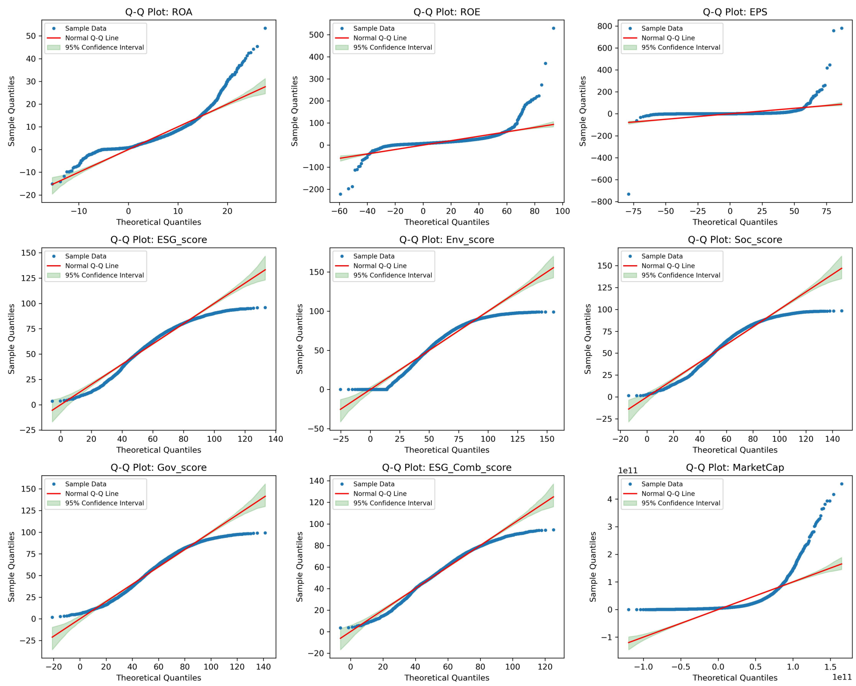


Figure 8.3: Q-Q Plot with 95% Confidence Interval for Assessing Normality for the European companies.

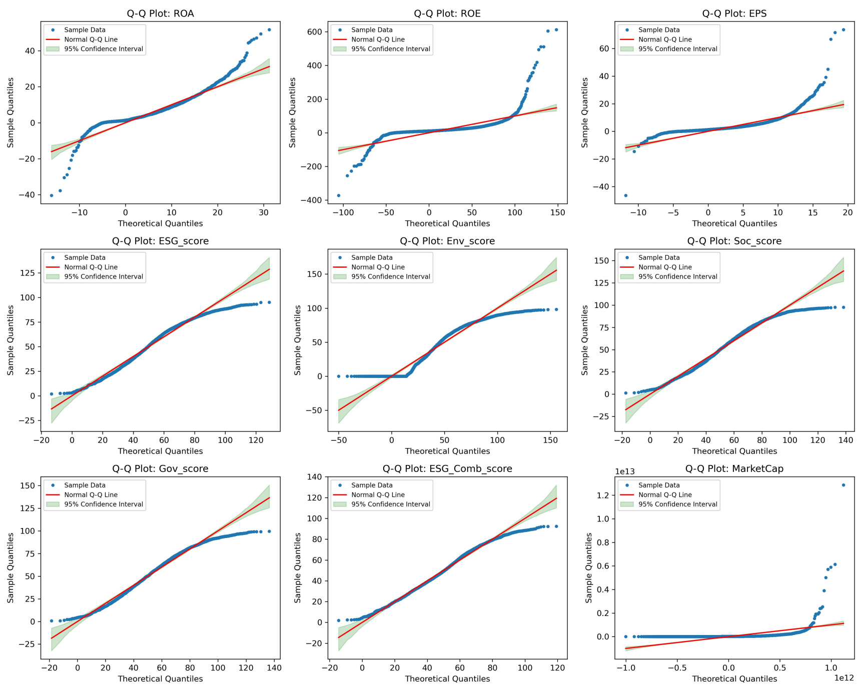


Figure 8.4: Q-Q Plot with 95% Confidence Interval for Assessing Normality for the USA companies.

Figure 8.5: Scatter Plot with Linear Regression and R^2 of all ESG variables vs Financial ratios for European dataset.

Figure 8.6: Scatter Plot with Linear Regression and R^2 of all ESG variables vs Financial ratios for American dataset.

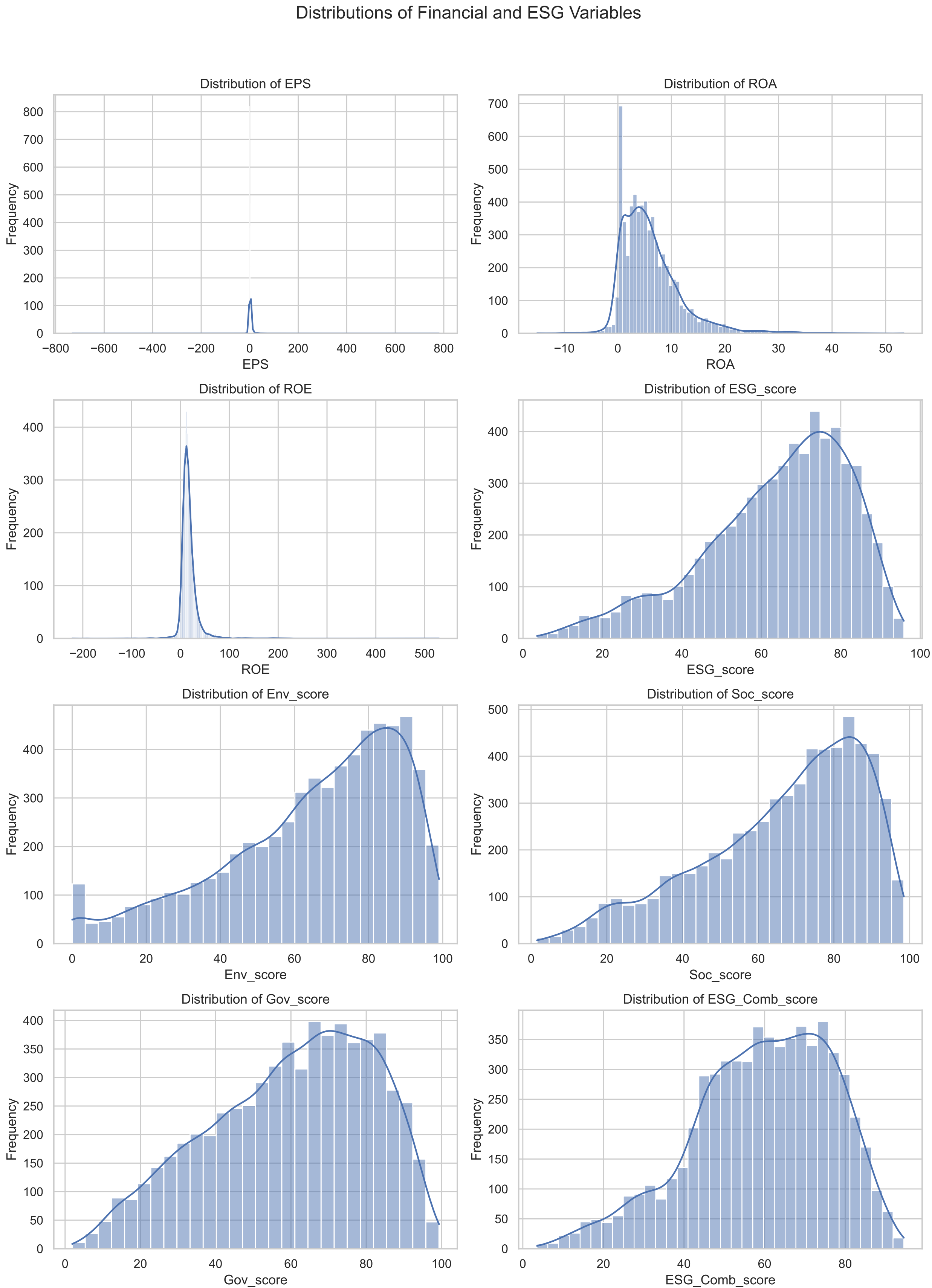


Figure 8.7: Distributions of Financial and ESG Variables for European data.

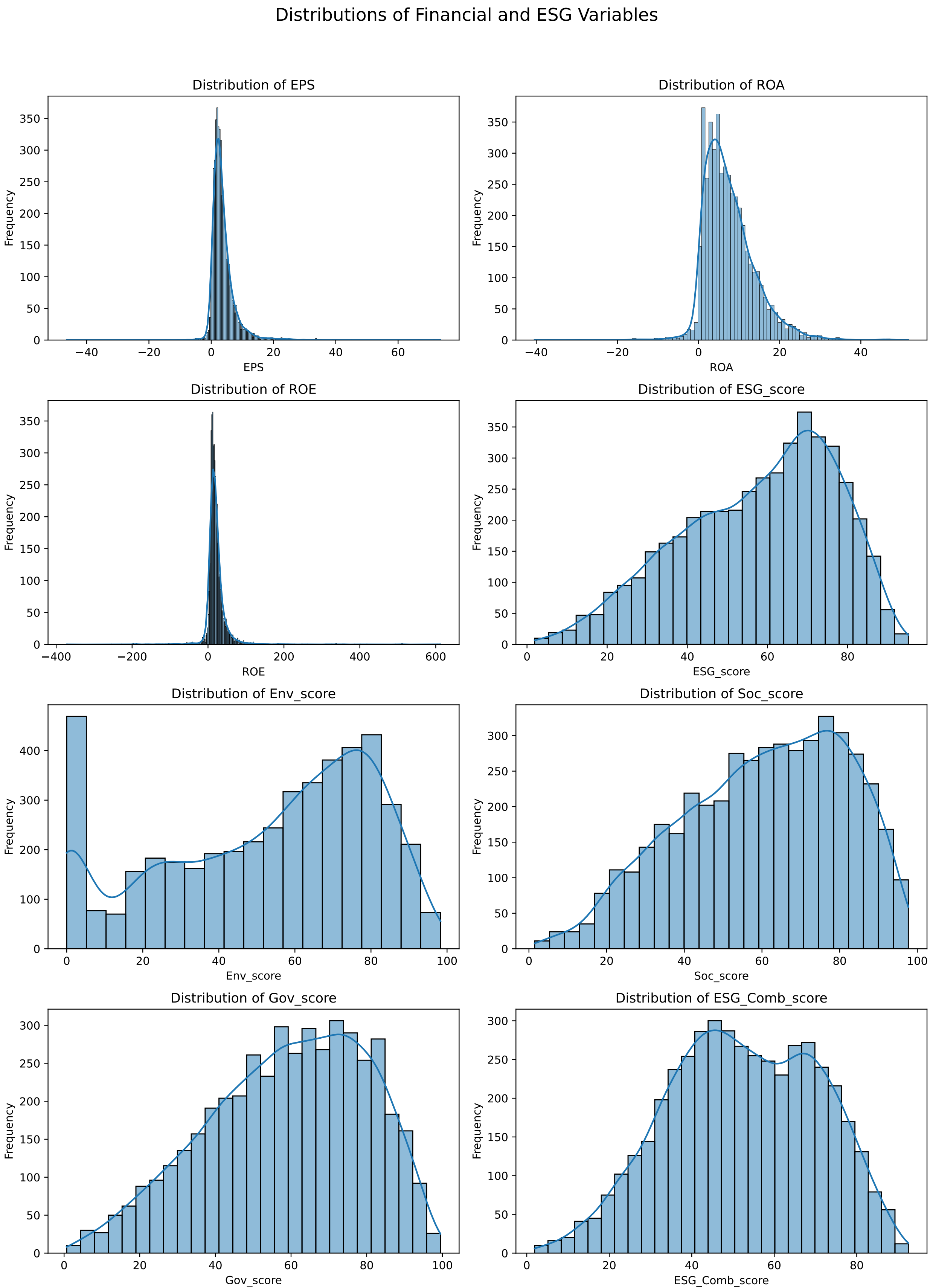


Figure 8.8: Distributions of Financial and ESG Variables for American data.

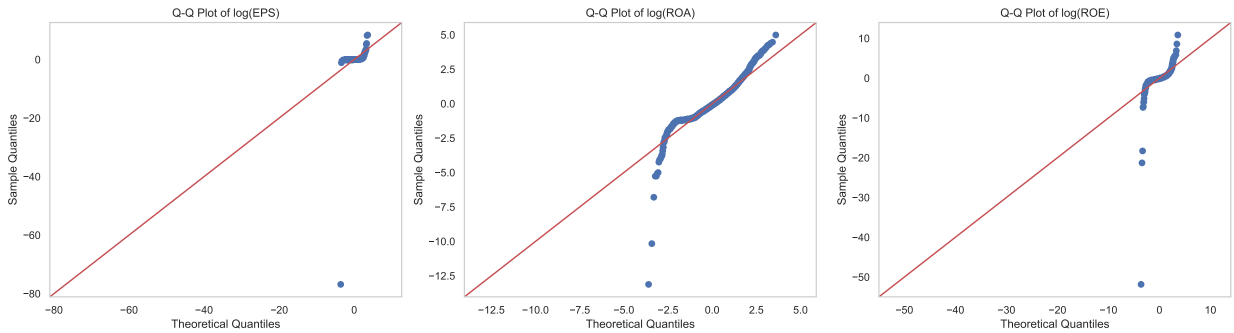


Figure 8.9: Q-Q plots for financial ratios after Log transformation for European data.

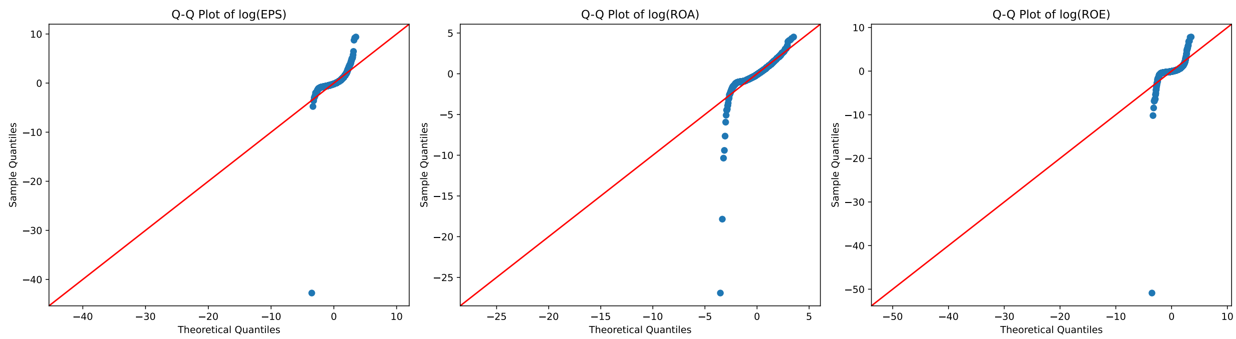


Figure 8.10: Q-Q plots for financial ratios after Log transformation for American data.

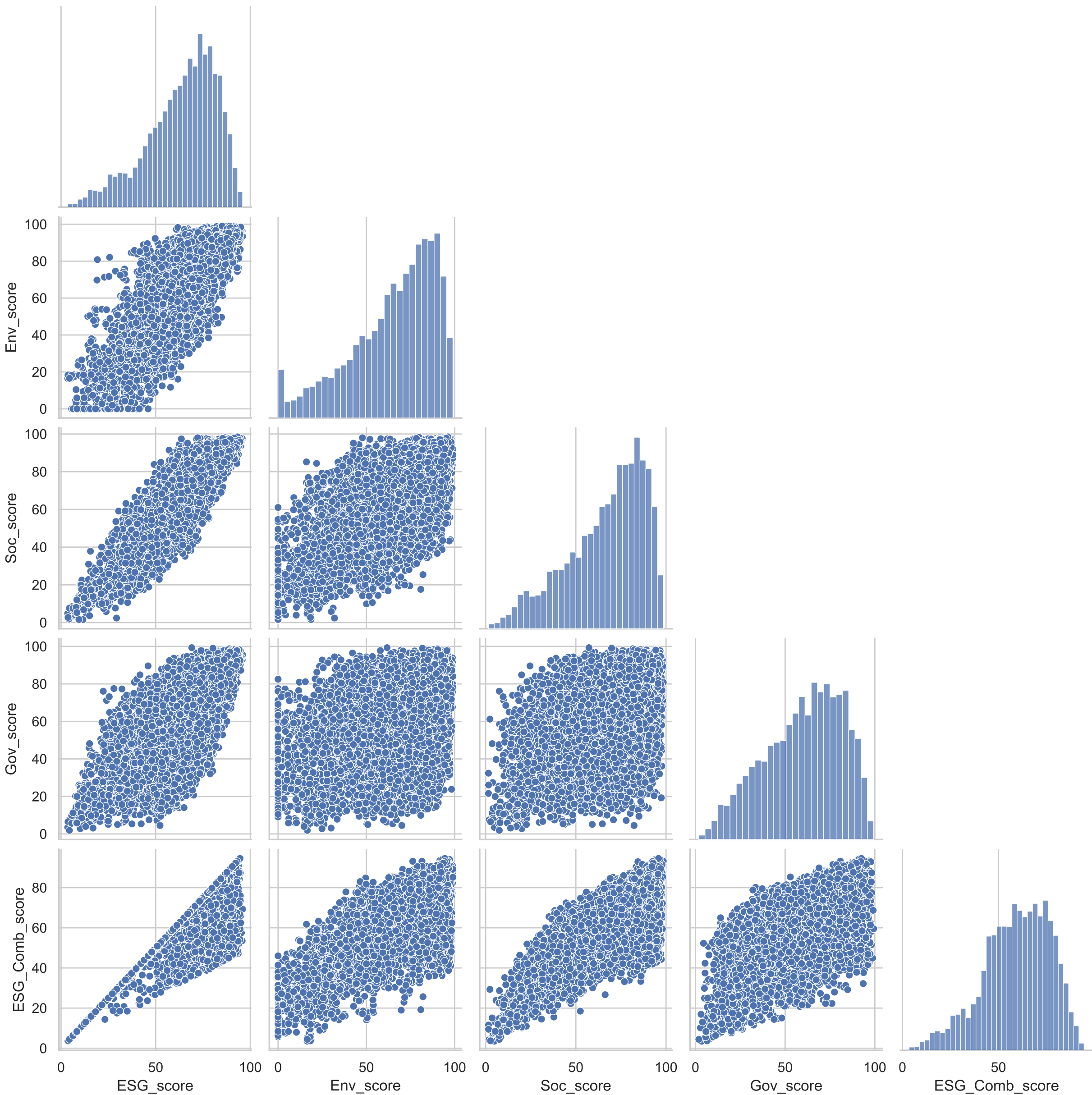


Figure 8.11: Scatter plot of ESG variables for European data.

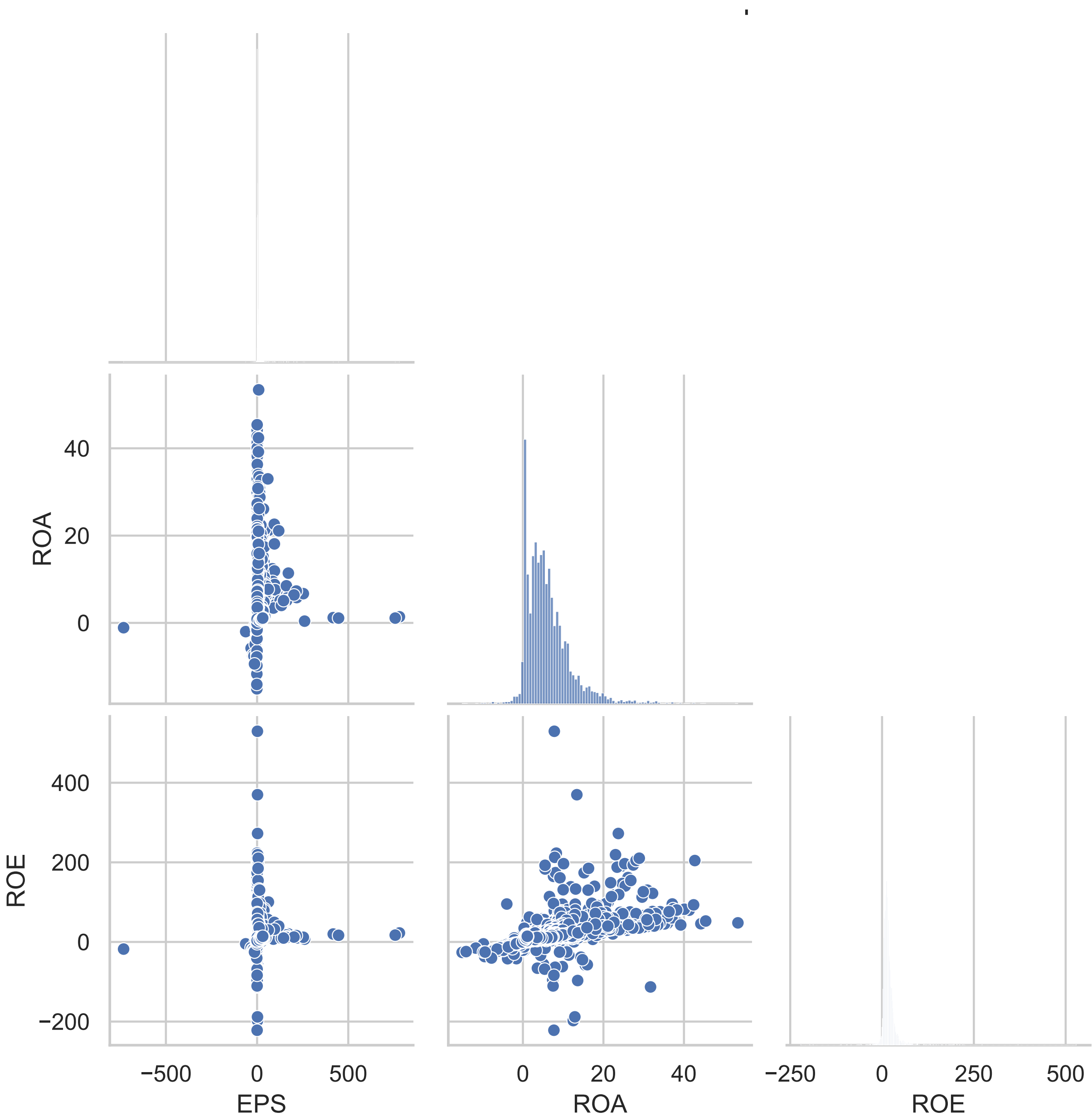


Figure 8.12: Scatter plot of financial ratios for European data.

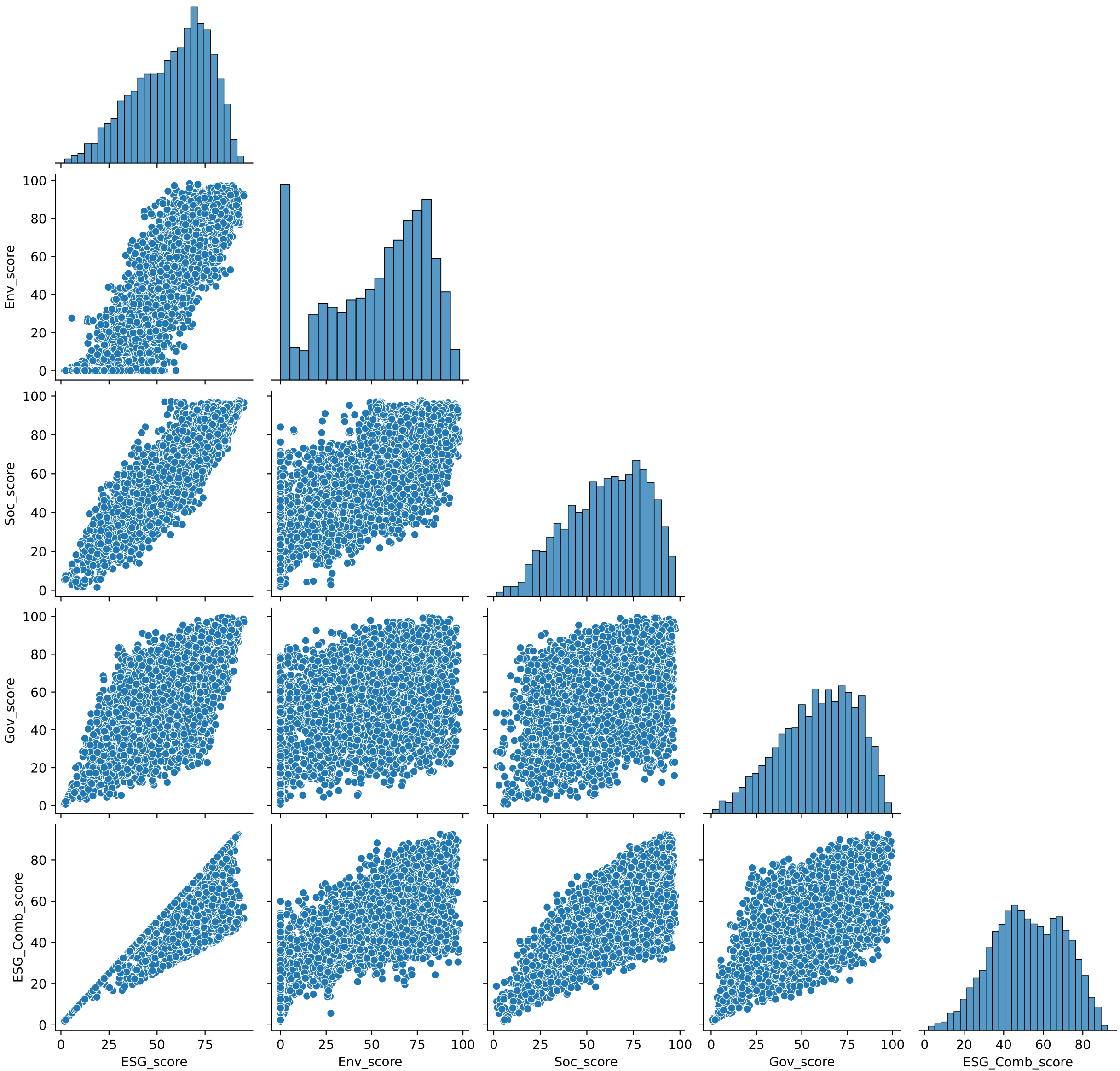


Figure 8.13: Scatter plot of ESG variables for American data.

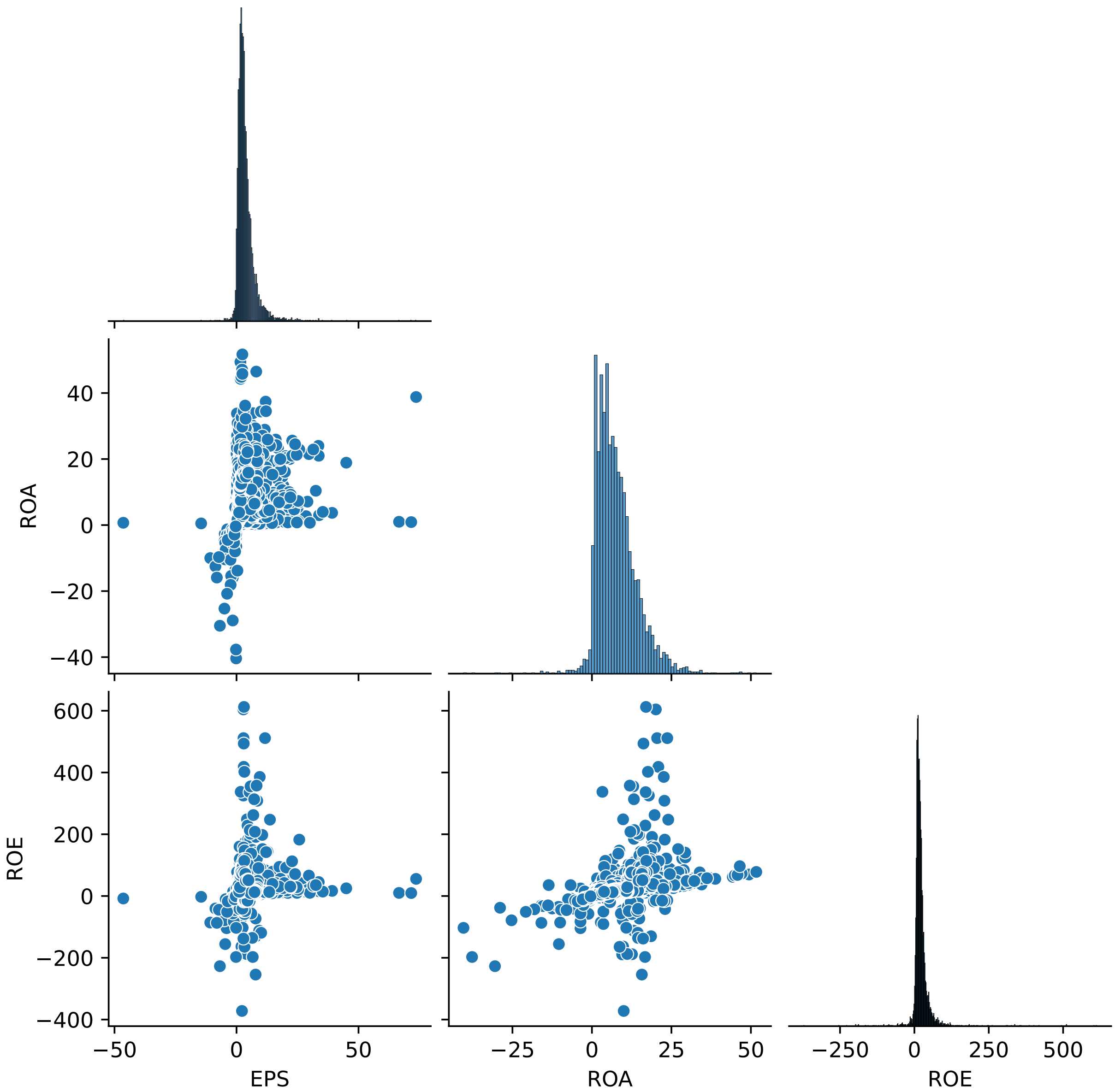


Figure 8.14: Scatter plot of financial ratios for American data.

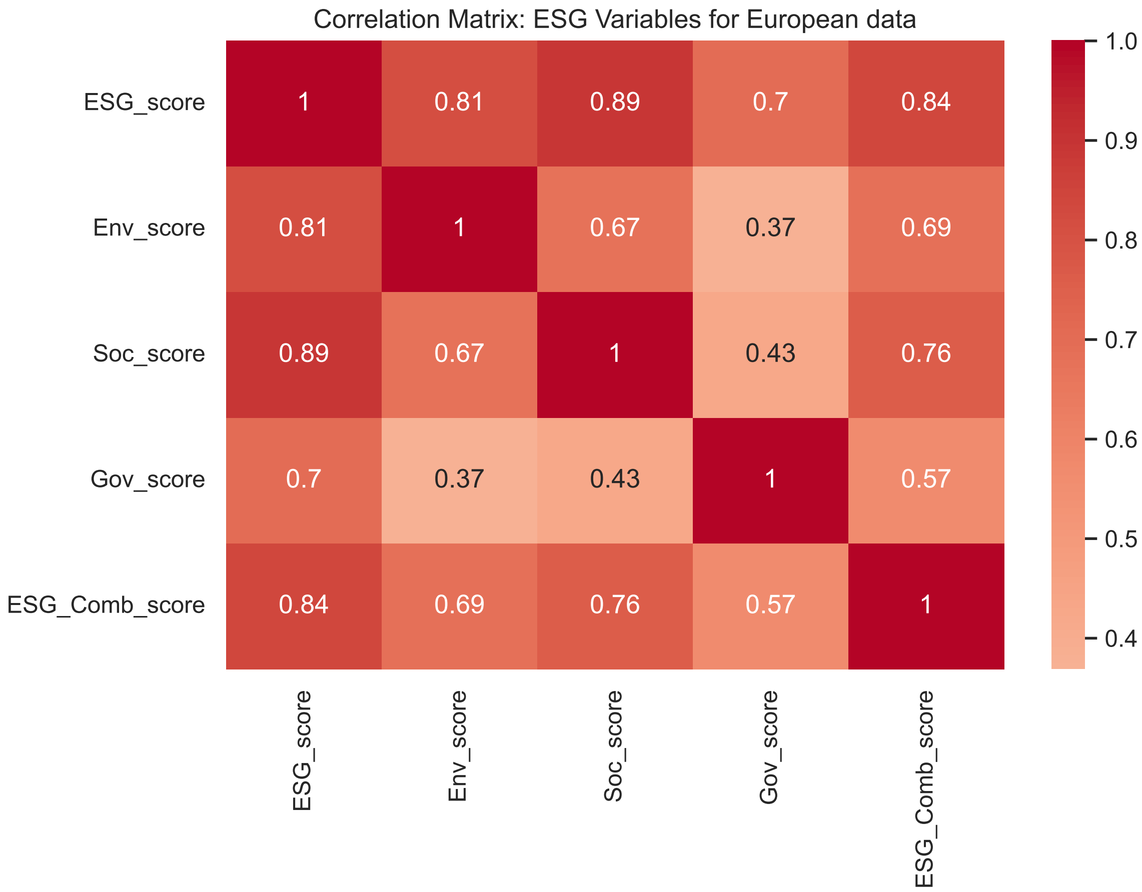


Figure 8.15: Correlation matrix of ESG variables for European data.

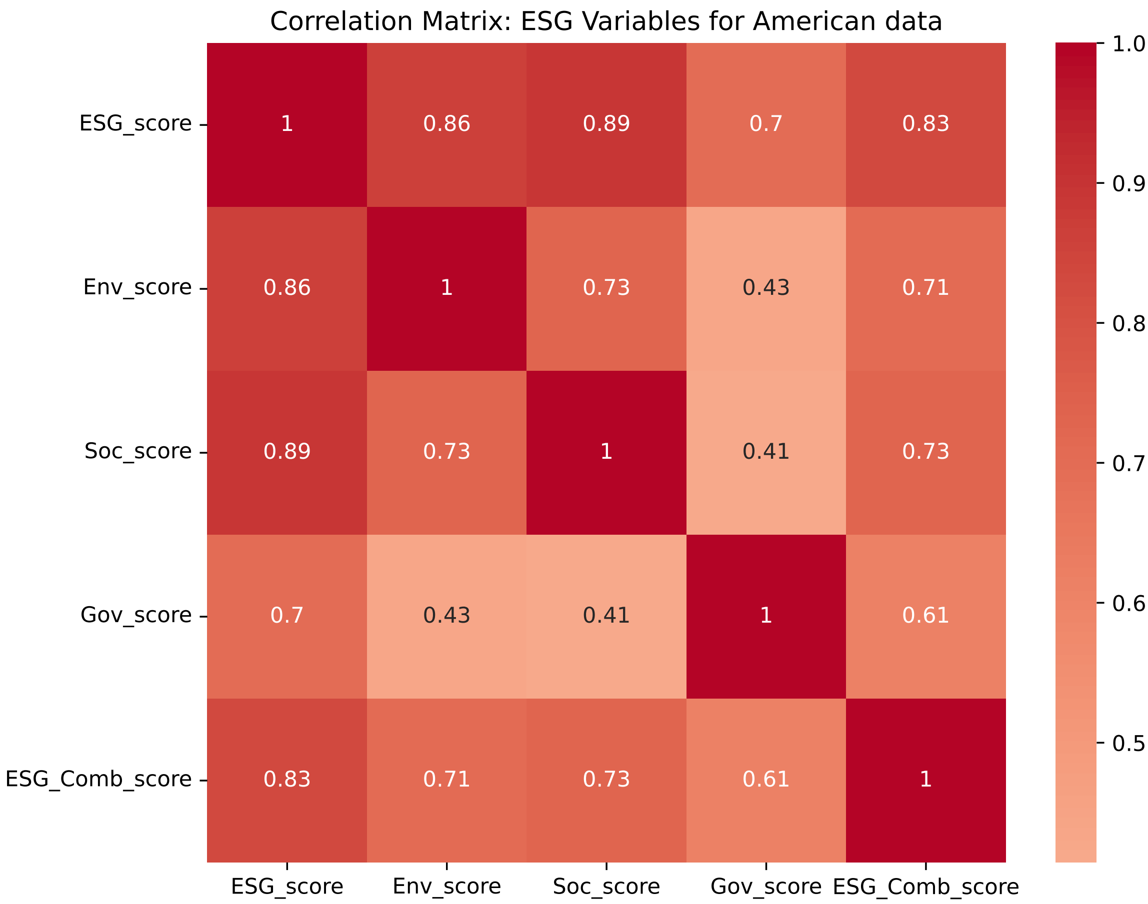


Figure 8.16: Correlation matrix of ESG variables for American data.

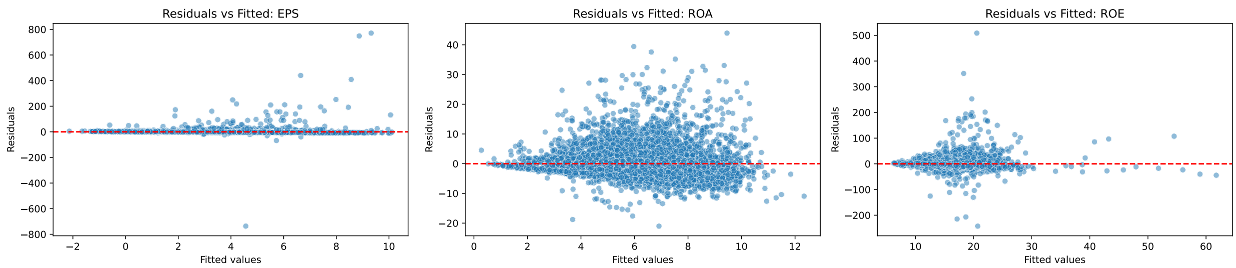


Figure 8.17: Residuals vs Fitted values for target values financial ratios for European data.

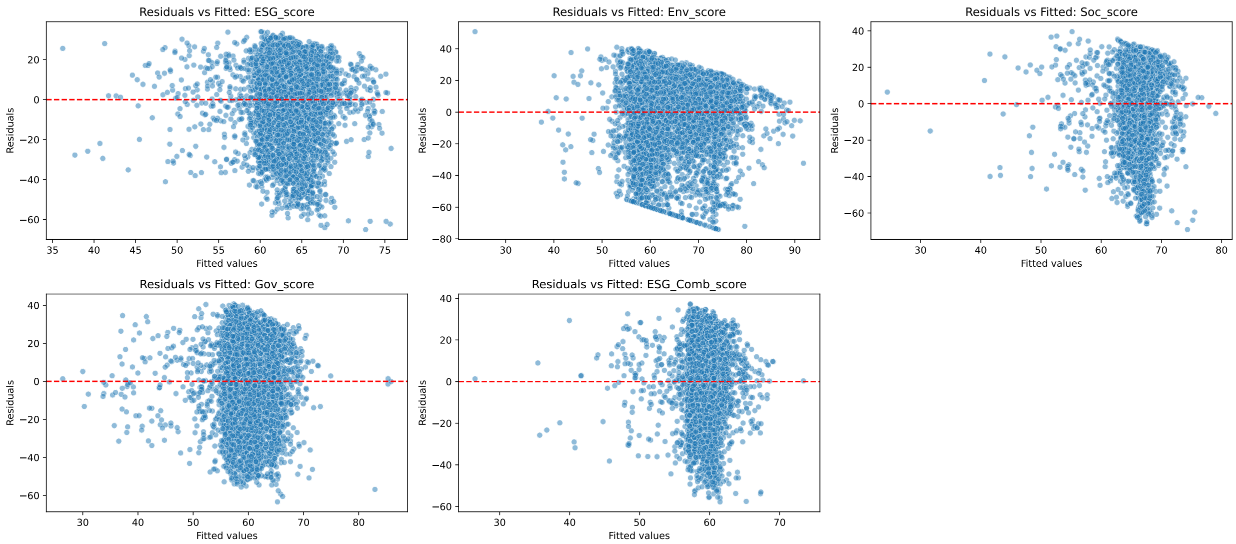


Figure 8.18: Residuals vs Fitted values for target values ESG variables for European data.

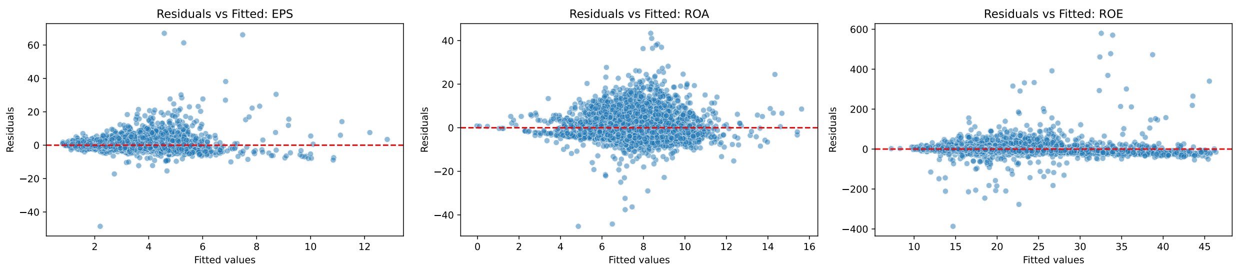


Figure 8.19: Residuals vs Fitted values for target values financial ratios for American data.

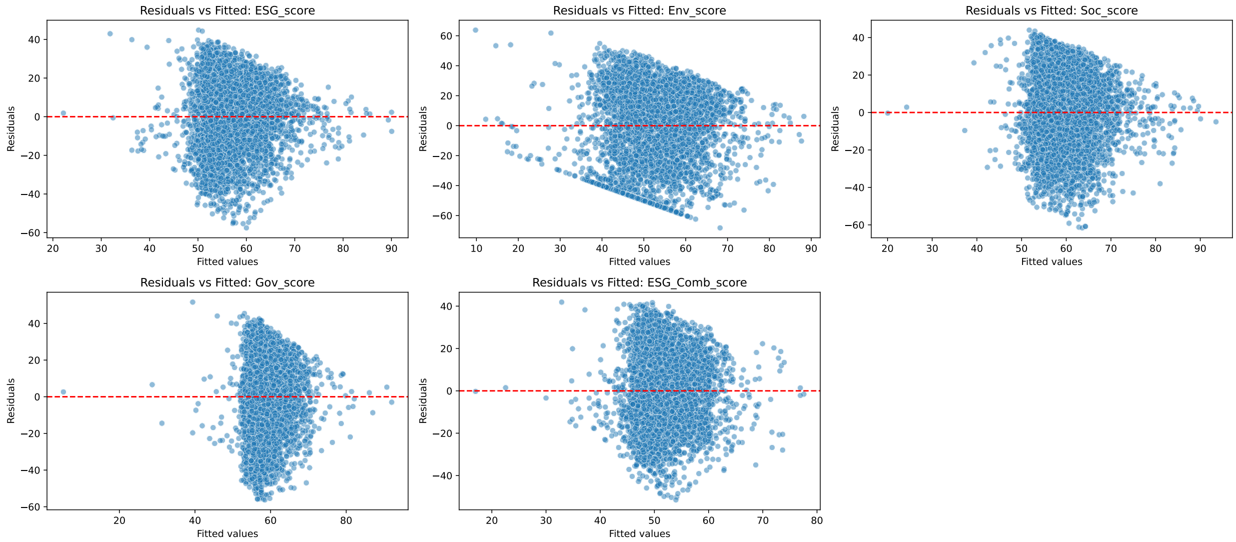


Figure 8.20: Residuals vs Fitted values for target values ESG variables for American data.

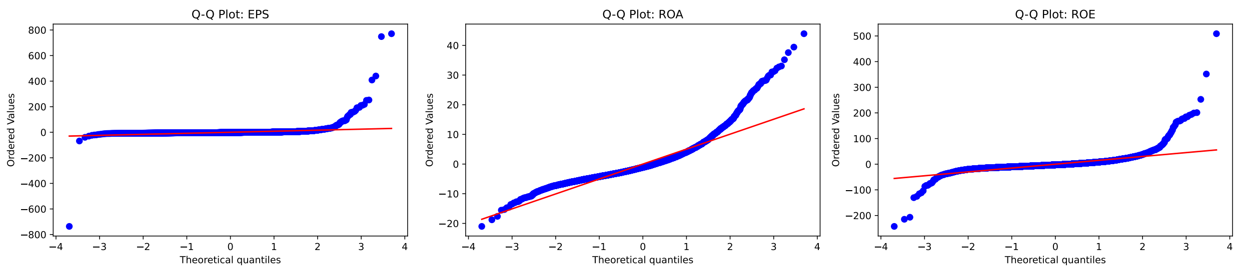


Figure 8.21: Q-Q plots for Normallity check for residuals with target values financial ratios for European data.

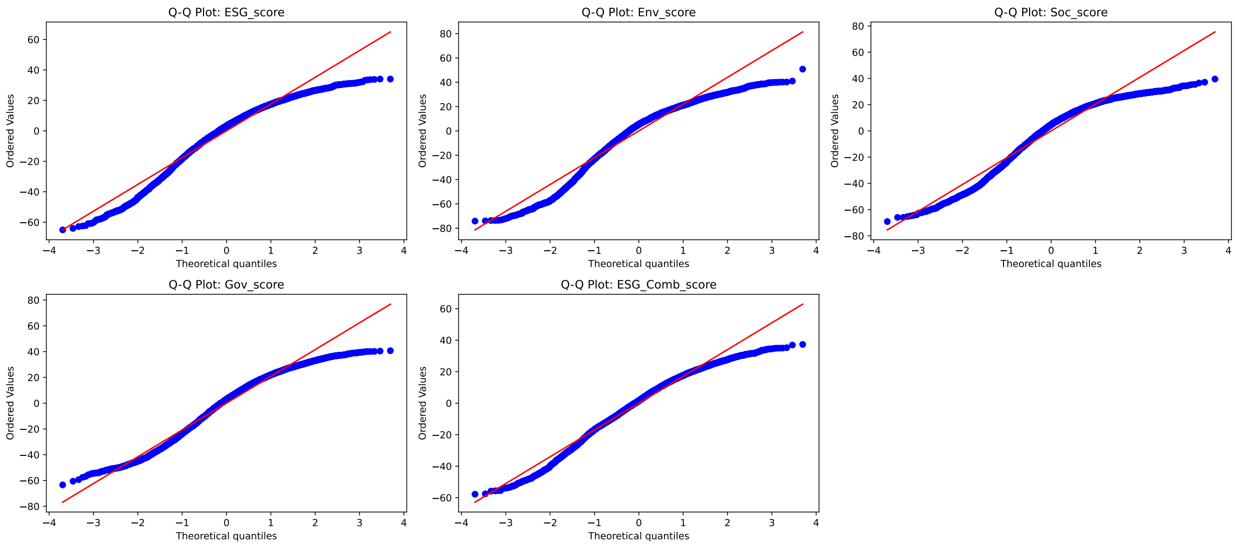


Figure 8.22: Q-Q plots for Normallity check for residuals with target values ESG variables for European data.

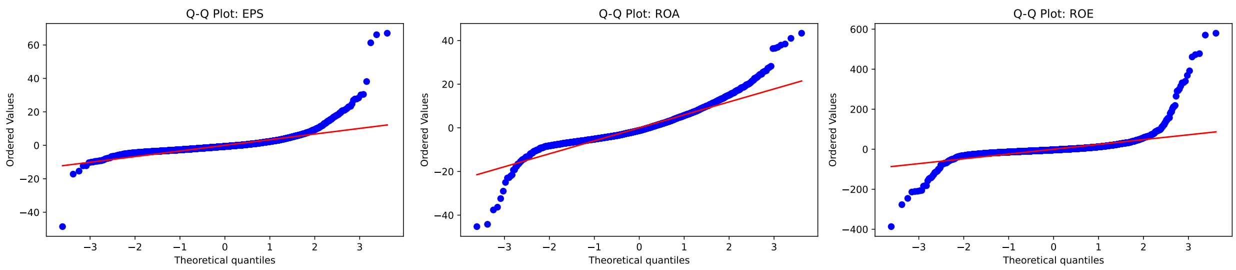


Figure 8.23: Q-Q plots for Normallity check for residuals with target values financial ratios for American data.

Industry	Companies
Commercial Banking	39
Direct Life Insurance Carriers	13
Pharmaceutical Preparation Manufacturing	13
Lessors of Nonresidential Buildings	13
Portfolio Management and Investment Advice	11
Wireless Telecommunications Carriers	11
Automobile and Light Duty Motor Vehicle Manufacturing	8
Lessors of Residential Buildings and Dwellings	8
Direct Property and Casualty Insurance Carriers	7
Electromedical and Electrotherapeutic Apparatus Manufacturing	6

(a) Top 10 Industries in Europe

Industry	Companies
Lessors of Nonresidential Buildings	11
Commercial Banking	10
Software Publishers	7
Pharmaceutical Preparation Manufacturing	7
Portfolio Management and Investment Advice	7
Direct Property and Casualty Insurance Carriers	6
Crude Petroleum Extraction	6
Electromedical and Electrotherapeutic Apparatus Manufacturing	6
Investment Banking and Securities Intermediation	5
Petroleum Refineries	5

(b) Top 10 Industries in USA

Table 8.2: Comparison of the top 10 industries in (a) Europe and (b) the USA based on the number of companies.

Company Name	**ESG-Score**									
	Dates					**ESG Scores**				
	FY0	**FY1**	**FY2**	**FY3**	**FY4**	**FY0**	**FY1**	**FY2**	**FY3**	**FY4**
Omv AG	31-12-2023	31-12-2022	31-12-2021	31-12-2020	31-12-2019	83.85	82.30	83.87	81.26	77.94
Verbund AG	31-12-2023	31-12-2022	31-12-2021	31-12-2020	31-12-2019	62.66	69.24	74.07	73.82	74.32
voestalpine AG	31-03-2024	31-03-2023	31-03-2022	31-03-2021	31-03-2020	70.77	57.80	64.58	64.83	70.14
Erste Group Bank AG	31-12-2023	31-12-2022	31-12-2021	31-12-2020	31-12-2019	71.76	76.08	78.37	80.61	84.11
Umicore SA	31-12-2023	31-12-2022	31-12-2021	31-12-2020	31-12-2019	78.10	81.20	75.11	73.71	75.13
Ackermans & Van Haaren NV	31-12-2023	31-12-2022	31-12-2021	31-12-2020	31-12-2019	51.37	44.33	40.18	37.79	44.27
Cofinimmo SA	31-12-2022	31-12-2021	31-12-2020	31-12-2019	31-12-2018	89.24	82.90	85.19	81.52	78.01
D'Ieteren Group SA	31-12-2023	31-12-2022	31-12-2021	31-12-2019	31-12-2018	78.75	75.86	73.55	75.97	62.17
Kbc Groep NV	31-12-2023	31-12-2022	31-12-2021	31-12-2020	31-12-2019	62.04	59.51	61.55	62.50	66.89
Ucb SA	31-12-2023	31-12-2022	31-12-2021	31-12-2020	31-12-2019	81.28	88.23	82.13	84.58	82.08
Ambu A/S	30-09-2023	30-09-2022	30-09-2021	30-09-2020	30-09-2019	71.78	68.79	59.78	55.79	54.20
Carlsberg A/S	31-12-2023	31-12-2022	31-12-2021	31-12-2020	31-12-2019	73.57	70.53	72.61	75.21	73.49
Coloplast A/S	30-09-2023	30-09-2022	30-09-2021	30-09-2020	30-09-2019	72.52	79.65	78.34	74.02	62.35
AP Moeller - Maersk A/S	31-12-2023	31-12-2022	31-12-2021	31-12-2020	31-12-2019	63.54	67.50	64.39	64.67	66.60

Table 8.3: Example of Company ESG Data for five financial years (FY0–FY4).

Company Name	2024	2023	2022	2021	2020	2019	2018
Omv AG	–	83.85	82.30	83.87	81.26	77.94	77.00
Verbund AG	–	62.66	69.24	74.07	73.82	74.32	68.01
voestalpine AG	70.77	57.80	64.58	64.83	70.14	61.68	61.35
Erste Group Bank AG	–	71.76	76.08	78.37	80.61	84.11	83.17
Umicore SA	–	78.10	81.20	75.11	73.71	75.13	74.87
Ackermans & Van Haaren NV	–	51.37	44.33	40.18	37.79	44.27	41.04
Cofinimmo SA	–	–	89.24	82.90	85.19	81.52	78.01
D'Ieteren Group SA	–	78.75	75.86	73.55	75.97	62.17	50.93
Kbc Groep NV	–	62.04	59.51	61.55	62.50	66.89	65.04
Ucb SA	–	81.28	88.23	82.13	84.58	82.08	80.35
Ambu A/S	–	71.78	68.79	59.78	55.79	54.20	51.05
Carlsberg A/S	–	73.57	70.53	72.61	75.21	73.49	69.55
Coloplast A/S	–	72.52	79.65	78.34	74.02	62.35	60.37
AP Moeller - Maersk A/S	–	63.54	67.50	64.39	64.67	66.60	61.76

Table 8.4: Company ESG Scores (2018–2024).

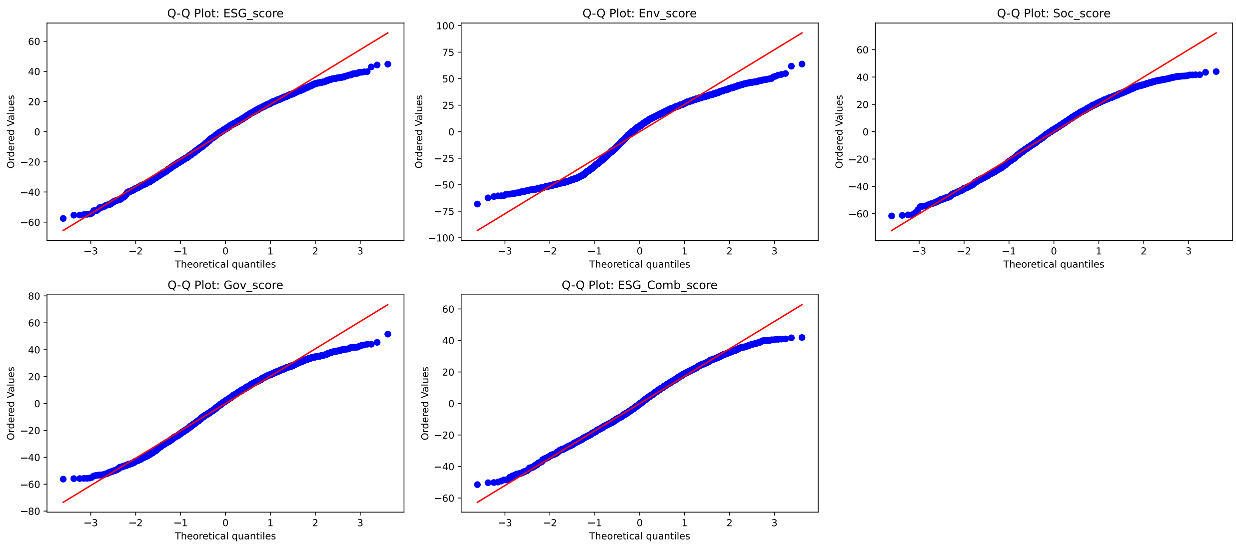


Figure 8.24: Q-Q plots for Normallity check for residuals with target values ESG variables for American data.

Company Name	Year	ESG Score
Omv AG	2024	–
Omv AG	2023	83.85
Omv AG	2022	82.30
Omv AG	2021	83.87
Omv AG	2020	81.26
Omv AG	2019	77.94
Omv AG	2018	77.00
Verbund AG	2024	–
Verbund AG	2023	62.66
Verbund AG	2022	69.24
Verbund AG	2021	74.07
Verbund AG	2020	73.82
Verbund AG	2019	74.32
Verbund AG	2018	68.01
voestalpine AG	2024	70.77
voestalpine AG	2023	57.80
voestalpine AG	2022	64.58
voestalpine AG	2021	64.83
voestalpine AG	2020	70.14
voestalpine AG	2019	61.68
voestalpine AG	2018	61.35
Erste Group Bank AG	2024	–
Erste Group Bank AG	2023	71.76
Erste Group Bank AG	2022	76.08
Erste Group Bank AG	2021	78.37
Erste Group Bank AG	2020	80.61
Erste Group Bank AG	2019	84.11
Erste Group Bank AG	2018	83.17
Umicore SA	2024	–
Umicore SA	2023	78.10
Umicore SA	2022	81.20
Umicore SA	2021	75.11
Umicore SA	2020	73.71
Umicore SA	2019	75.13
Umicore SA	2018	74.87

Table 8.5: Company ESG Scores for the first four companies (Long Format).

Rank	EPS	ROE	ROA	Env	Soc	Gov
1	409.63	56.02	32.15	25.91	8.47	8.44
2	571.12	50.49	42.41	53.07	17.68	2.02
3	684.52	200.24	17.75	1.02	97.31	55.67
4	718.91	167.46	0.00	95.33	98.06	44.64
5	261.09	53.90	31.70	19.51	17.17	11.29
6	697.08	192.90	16.63	9.84	99.99	61.21
7	780.29	34.82	20.15	81.42	66.97	62.75
8	780.29	409.89	2.16	31.30	65.54	9.01
9	780.29	73.17	10.65	74.56	23.97	0.00
10	780.29	91.84	1.57	96.69	47.62	13.87

Table 8.6: Top 10 Pareto Solutions – Europe

Rank	EPS	ROE	ROA	Env	Soc	Gov
1	73.57	233.53	0.00	99.24	62.39	13.28
2	28.21	0.64	33.31	3.15	3.29	2.43
3	40.18	0.20	33.46	3.55	8.66	2.67
4	25.01	387.92	11.43	7.68	99.32	29.31
5	36.62	23.34	34.32	3.58	5.98	0.01
6	73.31	232.32	0.41	100.00	61.94	11.53
7	21.97	145.87	39.07	12.71	99.79	16.63
8	70.83	218.78	14.86	73.70	100.00	23.00
9	64.31	0.64	32.52	0.01	14.49	2.20
10	18.69	172.88	27.16	24.48	95.52	11.42

Table 8.7: Top 10 Pareto Solutions – USA