



UNIVERSITÀ DI PAVIA

DEPARTMENT OF ECONOMICS AND MANAGEMENT

MASTERS PROGRAMME IN MSc FINANCE

TITLE: Relationship Between ESG & Company Financials

TITOLO: Relazione tra ESG e performance finanziaria delle imprese

Supervisor: Dr. Paolo Giudici

Prof: Dr. Paolo Giudici

Student:

Gurpreet Kaur Gurpreet Kaur

Academic Year 2025-2026

INDEX

Sr. No.	Title	Page No.
1.	Bibliography/ References	
2.	Chapter 1. Introduction 1.1 Background of ESG 1.2 Research Problem 1.3 Research Objective 1.4 Research questions 1.5 Research Contribution 1.6 Thesis Structure	7- 17
3.	Chapter 2. Literature Review 2.1 Evolution of ESG 2.2 Theoretical Foundation 2.3 ESG and Financial Performance literature 2.4 Determinants of ESG Performance 2.5 ML in ESG Research 2.6 Research Gap	18 - 29
3.	Chapter 3. Methodology 3.1 Research Design 3.2 Data Description 3.3 Variable definition 3.4 Econometric Models 3.5 Machine learning models 3.6 Evaluation Metrics	30 - 40
4.	Chapter 4. Empirical Results 4.1 Introduction 4.2 Initial ESG – Firm performance Analysis (1-year obser.) 4.3 Extended ESG – Firm Panel Analysis 4.4 Financial Performance as a determinant of ESG Performance	41 - 84
5.	Chapter 5. Discussion 5.1 Introduction 5.2 Discussion of Financial performance & ESG Relationship 5.3 Comparison with Previous Literature	85 - 88

	5.4 Discussion of Panel Data Finding 5.5 Discussion of Machine Learning Finding 5.6 Overall Interpretation of the Revised Framework	
6.	Chapter 6. Conclusion 6.1 Introduction 6.2 Summary of Major findings 6.3 Contribution of Study 6.4 Practical Implication 6.5 limitation of study 6.6 Recommendation for Future Research 6.7 Final Conclusion	89 - 92

BIBLIOGRAPHY & REFERENCES

Pollman, E. (2022). *The Making and Meaning of ESG*. European Corporate Governance Institute (ECGI). Available

at: [ECGI Working Paper – The Making and Meaning of ESG](#)

Freeman, R. E. (1984). *Strategic Management: A Stakeholder Approach*. Boston: Pitman.

Friedman, M. (1970). *The Social Responsibility of Business is to Increase its Profits*. The New York Times Magazine.

Friede, G., Busch, T., & Bassen, A. (2015). ESG and Financial Performance: Aggregated Evidence from More than 2000 Empirical Studies. *Journal of Sustainable Finance & Investment*, 5(4), 210–233.

Tsang, A., Wang, K. T., Liu, S., & Yu, L. (2023). Integrating Corporate Social Responsibility Criteria into Executive Compensation and Firm Innovation: International Evidence. *Journal of Corporate Finance*.

Xie, J., Nozawa, W., Yagi, M., Fujii, H., & Managi, S. (2019). Do Environmental, Social, and Governance Activities Improve Corporate Financial Performance? *Business Strategy and the Environment*, 28(2), 286–300.

RIASSUNTO

La crescente attenzione verso le tematiche ambientali, sociali e di governance ha reso i fattori Environmental, Social and Governance (ESG) un elemento centrale nella valutazione della sostenibilità e della performance di lungo periodo delle imprese. La letteratura esistente ha analizzato prevalentemente l'effetto della performance ESG sui risultati finanziari aziendali. Tuttavia, un numero più limitato di studi ha esaminato la relazione inversa, ovvero il modo in cui la performance finanziaria possa influenzare o spiegare il livello di performance ESG. La presente tesi si inserisce in questo filone di ricerca e analizza il ruolo della solidità finanziaria come possibile determinante della performance ESG.

L'analisi empirica si basa su un campione di 30 grandi imprese appartenenti all'indice S&P 500, osservate nel tempo attraverso dati panel. La performance ESG rappresenta la variabile dipendente, mentre gli indicatori finanziari utilizzati come variabili esplicative includono il Return on Assets (ROA), il Return on Equity (ROE), il livello di indebitamento e il totale delle attività. Lo studio combina metodi econometrici tradizionali con tecniche di machine learning. In particolare, vengono applicati modelli di regressione lineare OLS, regressioni panel a effetti fissi, Random Forest, XGBoost e LASSO Regression. Tale approccio consente sia di valutare la significatività statistica delle relazioni sia di confrontare la capacità predittiva dei diversi modelli.

I risultati mostrano una relazione positiva e statisticamente significativa tra redditività finanziaria e performance ESG. In particolare, ROA e ROE emergono come i principali indicatori finanziari associati a livelli più elevati di ESG. Questo risultato suggerisce che le imprese più redditizie dispongono di maggiori risorse, capacità organizzative e flessibilità finanziaria per investire in iniziative ambientali, pratiche sociali, sistemi di governance e attività rivolte agli stakeholder. Al contrario, il livello di indebitamento e il totale delle attività mostrano un contributo esplicativo più limitato.

L'analisi basata sui modelli di machine learning conferma la rilevanza della redditività. Random Forest e XGBoost identificano ROE e ROA come le variabili più importanti nella previsione della performance ESG. Il modello LASSO rafforza ulteriormente questo risultato, mantenendo ROA e ROE come variabili principali e riducendo sensibilmente il ruolo delle variabili meno rilevanti, in particolare del totale delle attività. Nel confronto tra modelli, la regressione lineare e il LASSO presentano le migliori performance predittive, suggerendo che

la relazione tra indicatori finanziari ed ESG, nel campione analizzato, risulta relativamente stabile, interpretabile e non eccessivamente complessa.

La tesi contribuisce alla letteratura ESG adottando una prospettiva inversa rispetto agli studi tradizionali: invece di analizzare l'effetto dell'ESG sulla performance finanziaria, esamina il ruolo della performance finanziaria nella spiegazione dei risultati ESG. L'integrazione tra approcci econometrici e machine learning permette inoltre di rafforzare la robustezza dei risultati. Le evidenze ottenute offrono implicazioni utili per manager, investitori e decisori politici, evidenziando come la solidità finanziaria possa rappresentare una base importante per lo sviluppo di strategie di sostenibilità aziendale nel lungo periodo.

Chapter 1. Introduction

1.1 Background of ESG

Environmental, Social, and Governance (ESG) has become a major development in corporate finance and investment analysis. In the last twenty years, ESG has grown from a niche ethical idea to a widely used framework for investors, regulators, companies, and financial institutions to assess corporate sustainability and long-term value. Concerns about climate change, social responsibility, transparency, and stakeholder engagement have sped up the adoption of ESG principles worldwide. Today, ESG factors are often included in investment strategies, risk management, and corporate governance.

ESG has its roots in Corporate Social Responsibility (CSR) and Socially Responsible Investing (SRI), which focused on ethical business practices and the social effects of corporate actions. At first, investors mainly avoided companies in controversial industries like tobacco, gambling, and weapons. Over time, ESG shifted this approach to a broader framework that evaluates corporate sustainability along with financial performance. Pollman (2022) notes that ESG became a standard way for firms and investors to measure environmental responsibility, social accountability, and governance quality in financial decisions.

The term ESG gained international recognition following the publication of the United Nations Principles for Responsible Investment (UN PRI) in 2006. The UN PRI encouraged institutional investors to incorporate environmental, social, and governance considerations into investment analysis and ownership practices. Since then, ESG investing has expanded rapidly due to growing investor awareness regarding sustainability risks and long-term corporate resilience. Today, ESG performance is increasingly viewed not only as an ethical consideration but also as an indicator of operational quality, risk management efficiency, and strategic corporate stability.

ESG consists of three major dimensions: environmental, social, and governance performance. The environmental dimension evaluates how firms manage ecological risks and environmental sustainability issues, including carbon emissions, energy efficiency, waste management, and climate-related policies. The social dimension focuses on stakeholder relationships, employee welfare, labour rights, diversity and inclusion, customer protection, and community engagement. Governance refers to corporate leadership structures, board independence,

executive compensation, shareholder rights, transparency, and ethical business conduct. Together, these dimensions provide a multidimensional framework through which firms can be assessed beyond purely financial performance indicators.

The growing relevance of ESG has also been strongly connected to the expansion of sustainable finance. Sustainable finance refers to financial activities that integrate environmental and social considerations into investment and business decisions to promote long-term economic sustainability. In recent years, institutional investors, pension funds, banks, and asset managers have increasingly relied on ESG metrics to identify financially stable and socially responsible firms. According to contemporary sustainable finance literature, ESG integration enables investors to evaluate non-financial risks that may significantly affect firm value and long-term profitability. Consequently, ESG disclosure practices have become increasingly important in corporate reporting and investor communication.

From a theoretical perspective, ESG is closely associated with Stakeholder Theory, which was originally developed by Freeman (1984). Stakeholder Theory argues that firms should not operate solely for shareholder profit maximization, but should also consider the interests of employees, customers, suppliers, communities, and other stakeholders. Under this perspective, strong ESG performance reflects a firm's ability to maintain sustainable relationships with multiple stakeholder groups, thereby contributing to long-term organizational success. In contrast, Shareholder Theory, primarily associated with Friedman (1970), argues that the primary responsibility of firms is to maximize shareholder wealth. The debate between these two perspectives has significantly influenced ESG research and corporate governance discussions in recent years.

Another important theoretical framework explaining ESG adoption is Legitimacy Theory. According to this theory, firms engage in ESG disclosure and sustainability activities in order to gain social legitimacy and maintain public trust. Companies may voluntarily disclose ESG information to demonstrate compliance with societal expectations and regulatory standards. Similarly, Signalling Theory suggests that ESG disclosure acts as a positive signal to investors by indicating transparency, strong governance quality, reduced operational risk, and long-term strategic orientation.

Academic interest in ESG has increased substantially over the last decade. The growing volume of ESG-related research reflects the increasing importance of sustainability issues within corporate finance and investment literature. Previous studies have examined the relationship

between ESG performance and financial outcomes, including profitability, firm value, market performance, and investment efficiency. However, empirical findings remain mixed and inconclusive. Some studies report a positive relationship between ESG performance and financial performance, while others identify weak, insignificant, or even negative relationships.

One of the most influential contributions to ESG literature was provided by Friede, Busch, and Bassen (2015), who conducted a meta-analysis of more than 2,000 empirical studies examining the relationship between ESG and corporate financial performance. Their findings suggested that the majority of studies identified a positive association between ESG performance and financial outcomes. This result contributed significantly to the growing perception that ESG activities may enhance corporate reputation, operational efficiency, investor confidence, and long-term firm value.

Despite the rapid expansion of ESG research, important research gaps continue to exist. Most existing studies primarily examine ESG as an explanatory variable influencing financial performance indicators such as Return on Assets (ROA), Return on Equity (ROE), or market valuation. Comparatively fewer studies investigate the reverse relationship, where financial performance acts as a determinant of ESG performance. Furthermore, limited research has integrated traditional econometric approaches with modern machine learning models to analyse ESG determinants and predictive relationships simultaneously. Therefore, the present study contributes to the existing literature by reversing the conventional ESG-financial framework and examining how financial performance indicators influence ESG outcomes through both econometric and machine learning methodologies.

1.2 Research Problem

The relationship between Environmental, Social, and Governance (ESG) performance and corporate financial performance has become one of the most widely discussed topics in modern financial and sustainability research. Over the last decade, companies, investors, policymakers, and financial institutions have increasingly focused on ESG factors as indicators of long-term corporate sustainability and responsible business practices. As ESG reporting and sustainable investment activities continue to grow globally, researchers have attempted to determine whether ESG performance contributes positively to financial success.

A large portion of existing literature primarily examines ESG as an independent variable influencing financial performance indicators such as Return on Assets (ROA), Return on Equity (ROE), firm profitability, market value, and investment efficiency. Several studies suggest that firms with stronger ESG performance may benefit from improved reputation, lower operational risk, stronger investor confidence, and better long-term financial stability. Friede, Busch, and Bassen (2015), through a large-scale meta-analysis of more than 2,000 empirical studies, found that the majority of research reports a positive relationship between ESG activities and financial performance. However, despite the growing number of studies, empirical findings remain inconsistent and often contradictory.

While some researchers identify a significant positive relationship between ESG and financial performance, others report weak, insignificant, or even negative relationships. These inconsistencies create uncertainty regarding the true nature of the ESG-financial performance relationship. One major reason for these conflicting findings is that previous studies frequently differ in terms of sample size, geographical coverage, ESG measurement methods, industry focus, and econometric approaches. In addition, many studies continue to treat ESG performance only as a factor influencing financial outcomes, while comparatively limited attention has been given to the reverse relationship.

The present study identifies an important research gap in the existing literature. Instead of examining whether ESG affects financial performance, this research investigates whether financial performance indicators influence ESG performance. In other words, the study reverses the conventional analytical direction by treating ESG as the dependent variable and financial indicators such as ROA, ROE, debt, and total assets as explanatory variables. This reversed framework provides a different perspective for understanding whether financially stronger firms are more capable of investing in ESG-related activities and sustainability initiatives.

Another important limitation in prior literature is the relatively limited integration of machine learning techniques within ESG-financial relationship studies. Traditional econometric methods such as Ordinary Least Squares (OLS) regression and panel regression remain dominant in ESG research. Although these methods are effective for explaining linear relationships, they may not fully capture complex nonlinear interactions between financial and ESG variables. Recent advances in artificial intelligence and machine learning have introduced new opportunities for predictive financial analysis, variable importance evaluation, and

nonlinear modelling. However, limited studies have directly compared econometric models with machine learning approaches within the ESG context.

Therefore, this study addresses two major research gaps simultaneously. First, it contributes to the literature by reversing the traditional ESG-financial relationship framework and examining financial performance as a determinant of ESG performance. Second, it integrates multiple analytical approaches, including linear regression, panel data analysis, Random Forest, XGBoost, and LASSO regression, in order to compare predictive performance and identify the most influential financial determinants of ESG performance.

The findings of this research are expected to contribute both theoretically and practically. From a theoretical perspective, the study expands the existing ESG literature by introducing a reversed analytical framework supported by both econometric and machine learning methodologies. From a practical perspective, the results may help investors, policymakers, and corporate managers better understand how financial strength influences ESG engagement and sustainability performance across firms.

1.3 Research Objectives

The primary objective of this study is to examine the relationship between corporate financial performance and Environmental, Social, and Governance (ESG) performance by adopting a reversed analytical framework compared to most existing ESG studies. While previous research has commonly investigated the impact of ESG performance on financial outcomes, this study focuses on understanding whether financial strength and corporate financial indicators can explain variations in ESG performance across firms.

The study aims to analyse whether companies with stronger financial performance are more likely to demonstrate higher ESG performance. In order to achieve this objective, ESG is treated as the dependent variable, while financial indicators such as Return on Assets (ROA), Return on Equity (ROE), debt, and total assets are used as explanatory variables. This approach allows the study to explore whether profitability, financial stability, and firm size contribute to stronger sustainability performance and ESG engagement.

Another important objective of the research is to compare traditional econometric techniques with modern machine learning approaches in order to evaluate their effectiveness in predicting ESG performance. The study applies multiple analytical models, including linear regression,

panel data regression, Random Forest, XGBoost, and LASSO regression. Through this comparison, the research seeks to identify which methodology provides the strongest explanatory and predictive performance in the context of ESG analysis.

Furthermore, the study aims to identify the most influential financial determinants of ESG performance. By using variable importance techniques within machine learning models, the research evaluates which financial indicators contribute most significantly to ESG prediction. This objective is particularly important because it provides deeper insights into the relationship between corporate financial strength and sustainability engagement.

The study also seeks to contribute to the growing literature on ESG and sustainable finance by integrating econometric and machine learning methodologies within a single analytical framework. Since existing literature has only limitedly explored the reversed relationship between financial performance and ESG outcomes, this research attempts to provide additional empirical evidence and methodological contribution to the field.

Overall, the research objectives are designed to support a broader understanding of how financial performance may influence ESG behaviour and sustainability performance across firms over time.

1.3.1 Main Research Objective

The main objective of this study is to examine the impact of financial performance indicators on ESG performance by treating ESG as the dependent variable and financial variables as explanatory factors.

1.3.2 Specific Research Objectives

The study further aims to achieve the following specific objectives:

1. To analyse the relationship between financial performance indicators and ESG performance.
2. To evaluate the impact of Return on Assets (ROA) on ESG scores.
3. To examine the influence of Return on Equity (ROE) on ESG performance.
4. To investigate whether debt levels and total assets significantly affect ESG outcomes.

5. To compare the performance of traditional econometric models and machine learning models in predicting ESG performance.
6. To identify the most important financial variables influencing ESG performance using machine learning techniques.
7. To evaluate whether nonlinear machine learning models provide stronger predictive capability compared to traditional regression approaches.
8. To contribute to the existing ESG literature by introducing a reversed ESG-financial analytical framework.

1.3.3 Expected Contribution of the Study

This study is expected to contribute both academically and practically. Academically, the research extends existing ESG literature by focusing on the reverse relationship between financial performance and ESG outcomes, which remains relatively underexplored in previous studies. Methodologically, the study contributes by integrating traditional regression analysis with advanced machine learning techniques such as Random Forest, XGBoost, and LASSO regression.

From a practical perspective, the findings may help investors, financial analysts, and corporate managers better understand how corporate profitability and financial strength influence ESG engagement and sustainability performance. In addition, policymakers and regulatory institutions may benefit from understanding the financial conditions that support stronger ESG adoption among firms.

1.4 Research Questions

The relationship between Environmental, Social, and Governance (ESG) performance and corporate financial performance has attracted growing attention in both academic literature and practical financial analysis. Over the years, many researchers have attempted to determine whether ESG activities improve firm profitability, market value, operational efficiency, and long-term sustainability. As ESG investing continues to expand globally, understanding the connection between financial performance and ESG outcomes has become increasingly important for investors, corporations, and policymakers.

Most previous studies within ESG literature primarily examine ESG as an independent variable that influences financial performance. In other words, researchers have traditionally focused on whether firms with stronger ESG performance generate better profitability, stronger market valuation, or improved shareholder returns. Several studies report a positive relationship between ESG and financial performance, suggesting that sustainable business practices may improve corporate reputation, investor confidence, stakeholder relationships, and long-term value creation. However, other studies report mixed or insignificant findings, creating continuing debate regarding the true direction and strength of the ESG-financial relationship.

Although ESG literature has expanded significantly in recent years, relatively fewer studies examine the reverse relationship between financial performance and ESG outcomes. Limited attention has been given to understanding whether financially stronger firms are more capable of investing in sustainability initiatives and maintaining higher ESG standards. This creates an important research gap within existing literature. Companies with stronger profitability, higher returns, and greater financial stability may possess more resources to allocate toward environmental initiatives, social responsibility programs, governance improvements, and sustainability reporting practices. Therefore, examining financial performance as a determinant of ESG performance offers a different and potentially important perspective within ESG research.

The present study attempts to contribute to this underexplored area by reversing the conventional ESG-financial framework. Instead of treating ESG as the explanatory variable, this research considers ESG as the dependent variable and investigates whether financial indicators such as Return on Assets (ROA), Return on Equity (ROE), debt, and total assets significantly influence ESG performance. This reversed analytical direction represents one of the key contributions of the study.

In addition, another important limitation within existing ESG literature is the relatively limited use of machine learning techniques. Traditional econometric models such as Ordinary Least Squares (OLS) regression and panel data regression remain dominant in most ESG studies. While these models are effective for identifying linear relationships, they may not fully capture more complex or nonlinear interactions between financial variables and ESG outcomes. With the rapid development of artificial intelligence and predictive analytics, machine learning models are increasingly being applied within financial research because of their ability to

handle large datasets, detect nonlinear patterns, and evaluate variable importance more efficiently.

Therefore, this study also contributes methodologically by integrating traditional regression analysis with multiple machine learning techniques, including Random Forest, XGBoost, and LASSO regression. By comparing these models, the study seeks to evaluate whether machine learning approaches can provide stronger predictive performance and deeper insights into ESG determinants compared to traditional econometric methods.

Based on these research gaps and objectives, the study seeks to answer the following research questions:

Main Research Questions

1. Does corporate profitability significantly explain ESG performance?
2. Which financial variables have the strongest influence on ESG outcomes?
3. Do financial indicators such as ROA, ROE, debt, and total assets significantly affect ESG scores?
4. Are financially stronger firms more likely to demonstrate higher ESG performance?
5. Can machine learning models predict ESG performance more effectively than traditional regression models?
6. Which analytical model provides the best predictive accuracy for ESG analysis?
7. Does the reversed ESG-financial framework provide additional insights compared to traditional ESG research approaches?

Contribution of the Research Questions

The research questions developed in this study are important because they address both theoretical and methodological gaps within existing ESG literature. Theoretically, the study contributes by shifting the analytical direction of ESG research and investigating whether financial performance itself acts as a driver of ESG performance. Methodologically, the study contributes by combining econometric analysis with machine learning approaches in order to compare explanatory strength, predictive capability, and variable importance across different models.

The findings of this study may help investors, financial analysts, and corporate managers better understand the relationship between profitability and sustainability performance. In addition, the study may provide useful insights regarding how firms allocate financial resources toward ESG initiatives and whether stronger financial conditions contribute to improved sustainability engagement over time.

1.5 Thesis Structure

This thesis is divided into five main chapters. Each chapter is designed to gradually develop the theoretical foundation, methodological framework, empirical analysis, and overall conclusions of the study. The structure of the thesis is organized as follows:

Chapter 1 — Introduction

The first chapter introduces the general background of ESG and explains the growing importance of sustainability and responsible investing in modern financial markets. The chapter discusses the research problem, research objectives, and research questions that guide the study. It also explains the motivation behind reversing the traditional ESG-financial relationship by examining financial performance as a determinant of ESG performance. Finally, the chapter presents the overall structure of the thesis.

Chapter 2 — Literature Review and Theoretical Framework

The second chapter reviews the existing academic literature related to ESG and corporate financial performance. It discusses the historical evolution of ESG, sustainable finance, and the growing role of ESG reporting within global financial markets. The chapter further explains the major theoretical foundations supporting the study, including Stakeholder Theory, Shareholder Theory, Legitimacy Theory, and Signalling Theory. In addition, previous empirical findings on ESG and financial performance are critically reviewed in order to identify the research gaps addressed by the present study.

Chapter 3 — Research Methodology

The third chapter explains the research design and methodological framework adopted in the study. This chapter describes the dataset, sample selection process, variable definitions, and data preparation procedures. The study applies both econometric and machine learning approaches to analyse ESG performance. Traditional statistical models such as linear regression and panel regression are combined with machine learning models including Random Forest, XGBoost, and LASSO regression. The chapter also explains the evaluation metrics used to compare model performance, including RMSE, MAE, and R^2 values.

Chapter 4 — Empirical Analysis and Results

The fourth chapter presents the empirical findings obtained from the econometric and machine learning analyses. The chapter includes descriptive statistics, correlation analysis, regression outputs, panel data results, and machine learning model comparisons. Various graphs, plots, diagnostic tests, and variable importance analyses are interpreted individually in order to provide a detailed understanding of the relationship between financial indicators and ESG performance. The predictive performance of different analytical models is also compared and discussed.

Chapter 5 — Conclusion and Recommendations

The final chapter summarizes the major findings of the research and discusses their academic and practical implications. The chapter highlights the contribution of the study to ESG and sustainable finance literature, particularly through the reversed ESG-financial framework and the integration of machine learning methods. In addition, the chapter discusses the limitations of the study and provides recommendations for future research related to ESG analysis, financial performance, and predictive modelling.

Overall, the thesis structure is designed to provide a logical and systematic progression from theoretical discussion to empirical investigation, allowing the study to comprehensively examine how financial performance influences ESG outcomes across firms.

Chapter 2. Literature Review

The increasing importance of Environmental, Social, and Governance (ESG) practices within global financial markets has generated substantial academic and professional attention over recent decades. ESG has evolved from a relatively narrow concept associated with ethical investing into a globally recognized framework used to evaluate corporate sustainability, long-term risk exposure, and responsible business practices. Investors, regulators, corporations, and financial institutions increasingly integrate ESG considerations into investment decisions, corporate strategies, and risk management processes.

The growing relevance of ESG has also contributed to a significant expansion in academic research examining the relationship between ESG performance and corporate financial performance. Existing literature largely focuses on whether ESG activities improve firm profitability, market value, operational efficiency, and investor confidence. However, despite the rapid growth of ESG research, empirical findings remain mixed and often inconsistent across industries, countries, and methodological approaches.

Most prior studies primarily examine ESG as an explanatory variable affecting financial performance indicators such as Return on Assets (ROA), Return on Equity (ROE), firm value, and stock market returns. Comparatively fewer studies investigate the reverse relationship by examining whether financially stronger firms are more capable of improving ESG performance. This reversed analytical perspective represents an important research gap within existing ESG literature and forms the central motivation for the present study.

In addition, recent developments in artificial intelligence and predictive analytics have introduced new analytical opportunities within financial and sustainability research. Traditional econometric techniques such as linear regression and panel data models continue to dominate ESG literature; however, machine learning models are increasingly recognized for their ability to capture nonlinear relationships and improve predictive accuracy. Despite this growing interest, relatively limited studies compare traditional econometric methods with machine learning approaches in explaining ESG performance.

Therefore, this chapter reviews the major academic literature and theoretical foundations related to ESG and financial performance. The chapter begins with the evolution of ESG and sustainable finance, followed by the discussion of major theoretical frameworks supporting ESG research. Subsequently, prior empirical findings on ESG and financial performance

relationships are reviewed, including studies focusing on financial determinants of ESG performance. Finally, the chapter identifies the major research gaps addressed by the present study.

2.1 Evolution of ESG and Sustainable Finance

2.1.1 Origins of ESG

The foundations of ESG can be traced back to ethical investing and socially responsible investing (SRI), which emerged during the twentieth century. Early ethical investing strategies were largely influenced by religious and moral principles, where investors avoided industries considered socially harmful, including tobacco, gambling, alcohol, and weapons manufacturing. These approaches later evolved into broader discussions regarding corporate responsibility and sustainable business behaviour.

During the 1960s and 1970s, increasing public awareness of environmental pollution, labour rights, and corporate misconduct strengthened demands for greater corporate accountability. Environmental movements and sustainability discussions encouraged investors and policymakers to consider the wider social impact of corporate activities. As a result, the concept of Corporate Social Responsibility (CSR) became increasingly important within business and management literature.

CSR encouraged firms to voluntarily integrate social and environmental concerns into their operations. However, CSR frameworks often lacked standardized measurement systems and consistent reporting practices. Consequently, financial markets gradually shifted toward more structured approaches capable of evaluating sustainability performance in measurable terms, contributing to the development of modern ESG frameworks.

2.1.2 Development of Modern ESG

The modern ESG framework gained global recognition during the early 2000s. One of the most important milestones occurred in 2004 when the United Nations introduced the *Who Cares Wins* report, formally connecting ESG considerations with long-term financial performance and investment decision-making.

In 2006, the United Nations launched the Principles for Responsible Investment (PRI), encouraging institutional investors to integrate ESG factors into investment analysis and ownership practices. Since then, ESG investing has expanded rapidly across global capital markets.

Today, ESG ratings are widely produced by organizations such as MSCI, Refinitiv, and Bloomberg. Investors increasingly use ESG scores to evaluate corporate sustainability, governance quality, and long-term business resilience.

At the same time, governments and regulatory authorities have strengthened ESG disclosure requirements. Many firms now publish sustainability reports and ESG disclosures alongside traditional financial statements, reflecting the growing institutionalization of ESG within modern finance.

2.1.3 ESG and Sustainable Finance

The expansion of ESG has closely aligned with the broader development of sustainable finance. Sustainable finance refers to financial activities that incorporate environmental, social, and governance considerations into investment and business decisions while promoting long-term economic sustainability.

ESG investing has become one of the fastest-growing segments within global financial markets. Investors increasingly recognize that ESG-related risks, including climate change, governance failures, and reputational damage, may significantly affect firm performance and long-term financial stability. Consequently, ESG is increasingly viewed not only as an ethical concern but also as a strategic risk management tool.

Several studies suggest that firms with stronger ESG performance may benefit from lower financing costs, improved operational efficiency, enhanced reputation, and stronger investor confidence. However, empirical findings remain mixed across industries and methodologies, creating continuing debate within academic literature.

2.2 Theoretical Framework

2.2.1 Stakeholder Theory

Stakeholder Theory, developed by R. Edward Freeman (1984), argues that corporations should create value not only for shareholders but also for a broad range of stakeholders, including employees, customers, suppliers, communities, and society as a whole.

According to this theory, firms that effectively address stakeholder interests are more likely to achieve long-term sustainability and stronger organizational performance. ESG activities therefore represent an important mechanism through which firms can improve stakeholder relationships, corporate reputation, and operational stability.

Stakeholder Theory provides strong support for ESG investment and sustainability initiatives because it emphasizes long-term value creation rather than short-term profit maximization alone.

2.2.2 Shareholder Theory

In contrast, Shareholder Theory, associated with Milton Friedman (1970), argues that the primary objective of corporations is to maximize shareholder wealth. According to this perspective, corporate managers should focus primarily on profitability and financial returns rather than broader social or environmental objectives.

Supporters of Shareholder Theory often argue that ESG initiatives may increase operational costs and reduce short-term profitability. As a result, some researchers believe that extensive ESG investments may negatively affect firm financial performance, particularly in industries facing strong competitive pressures.

This theoretical perspective contributes to the ongoing academic debate regarding whether ESG activities enhance or reduce corporate financial performance.

2.2.3 Legitimacy Theory

Legitimacy Theory suggests that corporations seek social acceptance and legitimacy from society in order to maintain their operations and reputation. Firms therefore engage in ESG disclosure and sustainability reporting to demonstrate responsible behaviour and maintain positive public perception.

Under this theory, ESG reporting becomes a strategic tool through which companies attempt to align corporate activities with societal expectations. Firms facing environmental controversies or governance concerns may increase ESG disclosure in order to improve legitimacy and reduce reputational risk.

Legitimacy Theory is particularly relevant in explaining why firms increasingly invest in sustainability reporting and ESG transparency.

2.2.4 Signalling Theory

Signalling Theory argues that companies use ESG activities and disclosures to signal positive information regarding firm quality, governance strength, and long-term strategic orientation.

Strong ESG performance may signal lower operational risk, stronger management quality, and improved corporate transparency to investors and financial markets. As a result, firms with stronger ESG ratings may attract greater investor confidence and lower financing costs.

This theory is highly relevant to the present study because financially stronger firms may use ESG investment as a signal of long-term corporate stability and strategic capability.

2.3 ESG and Financial Performance Literature

2.3.1 Positive Relationship Between ESG and Financial Performance

A large proportion of existing literature reports a positive relationship between ESG performance and financial performance. Several studies argue that firms with strong ESG practices benefit from improved operational efficiency, stronger stakeholder relationships, reduced risk exposure, and enhanced corporate reputation.

One of the most influential studies in this area is conducted by Friede, Busch, and Bassen (2015), who reviewed more than 2,000 empirical studies and found that the majority reported a positive relationship between ESG and corporate financial performance.

Similarly, Xie et al. (2019) found that ESG activities may contribute to stronger profitability and improved firm value through enhanced operational performance and sustainability management.

Supporters of the positive ESG-financial relationship argue that sustainability initiatives improve long-term strategic positioning and increase investor confidence.

2.3.2 Negative Relationship Between ESG and Financial Performance

Despite the growing support for ESG, some studies report a negative relationship between ESG activities and financial performance. Critics argue that ESG implementation may increase operational costs, reduce managerial efficiency, and divert resources away from core business activities.

According to this perspective, firms investing heavily in ESG initiatives may experience lower short-term profitability due to increased sustainability expenditures, compliance costs, and reporting obligations.

Some researchers therefore argue that ESG investments may create financial burdens, particularly for firms operating in highly competitive industries with limited financial flexibility.

2.3.3 Mixed and Insignificant Findings

Several studies report mixed or statistically insignificant relationships between ESG and financial performance. Researchers argue that ESG outcomes may vary depending on industry characteristics, regional differences, firm size, methodology, and time horizon.

In some cases, environmental initiatives may improve long-term performance while creating short-term financial pressure. Similarly, governance quality may affect firms differently depending on institutional and regulatory environments.

The inconsistency of empirical findings demonstrates that the ESG-financial relationship remains complex and multidimensional. This ongoing debate highlights the importance of further research using alternative methodological approaches and analytical frameworks.

2.4 Financial Determinants of ESG Performance

2.4.1 Profitability and ESG Performance

Several studies suggest that financially stronger firms are more capable of investing in ESG initiatives due to greater financial flexibility and resource availability. This perspective is often supported by Slack Resource Theory, which argues that profitable firms possess excess resources that can be allocated toward sustainability investments.

Indicators such as Return on Assets (ROA) and Return on Equity (ROE) are frequently associated with stronger ESG performance because profitable firms may have greater capacity to invest in environmental programs, employee welfare, and governance improvements.

2.4.2 Firm Size and ESG Performance

Firm size is commonly identified as an important determinant of ESG performance. Larger firms generally face greater public scrutiny, regulatory pressure, and stakeholder expectations. Consequently, large corporations are often more active in ESG reporting and sustainability initiatives.

In addition, larger firms typically possess greater financial and managerial resources, allowing them to invest more extensively in ESG programs and sustainability disclosures.

2.4.3 Debt and ESG Performance

Leverage and debt levels may also influence ESG performance. Firms with high debt obligations may face financial constraints that limit their ability to invest in ESG activities. At the same time, creditors may increasingly evaluate ESG risks when making financing decisions.

Some studies suggest that firms with lower financial risk and more stable capital structures may demonstrate stronger ESG performance due to reduced financial pressure and improved long-term planning capacity.

2.5 Machine Learning and ESG Research

2.5.1 Artificial Intelligence in Finance

Artificial Intelligence (AI) has become increasingly important within modern financial markets and investment analysis. Financial institutions, investment firms, banks, and researchers now

use AI-based systems to process large volumes of financial data, improve decision-making, identify hidden patterns, and enhance forecasting accuracy. The rapid growth of computational power and data availability has significantly expanded the application of AI within financial research over the last decade.

AI techniques are particularly valuable in financial analysis because financial markets are highly dynamic, nonlinear, and influenced by multiple interconnected variables. Traditional statistical models often struggle to fully capture these complexities, especially when dealing with large datasets and rapidly changing market conditions. Consequently, researchers have increasingly adopted machine learning and AI-based methods to improve financial forecasting and predictive analysis.

Within sustainable finance, AI has become increasingly relevant for ESG analysis, sustainability assessment, corporate risk evaluation, and investment screening. ESG datasets often contain multidimensional information involving environmental indicators, governance quality, profitability measures, social performance, and firm-specific characteristics. AI methods allow researchers to analyse these complex relationships more effectively than traditional linear approaches alone.

As ESG investing continues to expand globally, the integration of AI into sustainability research has become an important development within both academic and professional financial analysis.

2.5.2 Predictive Analytics in ESG Research

Predictive analytics refers to the use of statistical models, machine learning techniques, and computational algorithms to forecast future outcomes based on historical data patterns. In ESG research, predictive analytics is increasingly used to estimate ESG performance, evaluate sustainability risks, identify corporate behaviour patterns, and forecast long-term financial outcomes.

The application of predictive analytics within ESG studies has grown substantially due to the increasing availability of corporate sustainability data and ESG ratings. Financial analysts and researchers now seek not only to explain ESG performance but also to predict future ESG behaviour using financial and non-financial indicators.

Predictive analytics offers several advantages within ESG research. First, it allows the identification of complex relationships between profitability, governance quality, leverage, firm size, and ESG outcomes. Second, predictive models can improve forecasting accuracy by detecting nonlinear interactions that may not be observable through conventional econometric methods. Third, predictive analytics allows comparative evaluation between different models in order to determine the most effective analytical approach.

As a result, predictive analytics has become an increasingly important component of modern ESG and sustainable finance research.

2.5.3 ESG Forecasting Through Machine Learning

Machine learning techniques are increasingly applied within ESG forecasting and sustainability analysis. Unlike traditional regression models, machine learning approaches can automatically identify patterns within large datasets while adapting to complex and nonlinear relationships between variables.

Several machine learning methods have become highly popular within ESG research, including Random Forest, XGBoost, Support Vector Machines, Neural Networks, and LASSO regression. These models are often used to:

- forecast ESG scores,
- predict sustainability performance,
- assess ESG-related financial risks,
- classify sustainable investments,
- and evaluate corporate governance quality.

Random Forest models are particularly useful for identifying variable importance and understanding which financial indicators contribute most strongly to ESG prediction. XGBoost models provide strong predictive accuracy and are highly effective in processing large financial datasets with complex interactions. LASSO regression is valuable for selecting the most important explanatory variables while reducing model complexity.

Machine learning models are especially relevant within ESG analysis because ESG performance is influenced by multiple interconnected financial and organizational factors.

Relationships between profitability, debt, governance, and sustainability performance may not always follow simple linear patterns. Consequently, machine learning methods may provide stronger predictive capabilities than traditional econometric techniques in certain situations.

The growing use of machine learning within ESG research demonstrates the increasing importance of computational methods within modern financial and sustainability analysis.

2.5.4 Machine Learning Versus Econometric Models

Traditional econometric models such as Ordinary Least Squares (OLS) regression and panel data regression remain widely used within ESG literature. These models provide strong interpretability and allow researchers to examine statistical significance, coefficient direction, and theoretical relationships between variables.

However, econometric models often rely on assumptions such as linearity, normality, and homoscedasticity. ESG and financial relationships may involve nonlinear interactions and hidden patterns that cannot always be fully captured through standard linear techniques.

Machine learning models differ from traditional econometric methods because they prioritize predictive performance and pattern recognition rather than solely focusing on statistical inference. Models such as Random Forest and XGBoost can process complex datasets more flexibly and may improve prediction accuracy when relationships between variables are highly nonlinear.

Nevertheless, machine learning methods also present limitations. Some advanced AI models may lack transparency and interpretability compared to regression-based approaches. As a result, combining econometric and machine learning methods may provide a more balanced analytical framework by integrating interpretability with predictive strength.

The comparison between machine learning and econometric models has therefore become an important topic within modern ESG research. Researchers increasingly seek to determine whether advanced machine learning techniques outperform traditional regression models in explaining ESG performance and sustainability outcomes.

2.6 Research Gap

Despite the rapid expansion of ESG literature, several important research gaps remain within existing academic studies.

First, the majority of prior research primarily examines ESG as an explanatory variable affecting financial performance. Most empirical studies focus on whether ESG activities improve profitability, operational efficiency, firm value, stock market returns, or investor confidence. As a result, the traditional research direction within ESG literature has largely concentrated on the relationship between ESG performance and financial outcomes.

Second, comparatively fewer studies examine the reverse relationship by analysing whether stronger financial performance contributes to higher ESG performance. Limited research investigates whether financially successful firms possess greater resources and capabilities to invest in sustainability initiatives, governance improvements, and social responsibility programs. Consequently, the financial determinants of ESG performance remain relatively underexplored within existing literature.

Third, although machine learning applications within finance have expanded significantly, relatively limited studies directly compare machine learning techniques with traditional econometric models in ESG analysis. Most ESG studies continue to rely heavily on regression-based approaches, while fewer studies evaluate whether machine learning models such as Random Forest, XGBoost, and LASSO regression provide stronger explanatory or predictive capabilities.

Fourth, limited studies combine both econometric and machine learning approaches within a single ESG framework. Comparative analysis between Linear Regression, Panel Regression, Random Forest, XGBoost, and LASSO regression remains relatively scarce, particularly within studies examining the financial determinants of ESG performance.

Finally, existing ESG literature often reports inconsistent findings regarding the relationship between sustainability and financial variables. Positive, negative, and mixed empirical results suggest that ESG-financial relationships may be more complex than traditional linear models can fully explain. This further strengthens the need for alternative analytical frameworks and predictive methodologies.

The present study attempts to address these research gaps by reversing the traditional ESG-financial framework and examining ESG performance as the dependent variable explained by

financial indicators such as Return on Assets (ROA), Return on Equity (ROE), Total Assets, and Debt. In addition, the study integrates traditional econometric techniques with multiple machine learning models in order to compare predictive performance and identify the strongest determinants of ESG performance.

Therefore, the study contributes to existing literature by:

- examining the reverse financial performance → ESG relationship,
- integrating econometric and machine learning approaches,
- comparing multiple predictive models,
- identifying key financial determinants of ESG performance,
- and providing additional evidence regarding the complexity of ESG-financial interactions.

Chapter 3. Research Methodology

3.1 Research Design

The present study adopts a quantitative research design in order to examine the relationship between financial performance and ESG performance among publicly listed companies. Unlike traditional ESG studies that primarily analyse the effect of ESG performance on financial outcomes, this research reverses the conventional framework by investigating whether financial indicators can explain variations in ESG performance.

The study combines both econometric and machine learning approaches in order to provide a broader analytical perspective. Traditional econometric models are used to examine statistical relationships and coefficient significance, while machine learning models are applied to evaluate predictive performance and identify nonlinear relationships between variables.

The research follows a deductive approach because the analysis is based on previously established theories regarding profitability, firm resources, corporate governance, and sustainability behaviour. Existing literature suggests that financially stronger firms may possess greater capabilities to invest in ESG initiatives, governance quality, environmental management, and social responsibility activities. Based on these theoretical assumptions, the study empirically evaluates whether financial indicators significantly influence ESG performance.

The overall research framework consists of:

- data collection and preprocessing,
- descriptive and correlation analysis,
- econometric regression analysis,
- panel data analysis,
- machine learning implementation,
- model comparison,

- and interpretation of predictive performance.

The study therefore integrates explanatory analysis with predictive analytics in order to provide a comprehensive understanding of the financial determinants of ESG performance.

3.2 Data Description

The dataset used in this study was collected from companies listed within the S&P 500 index. Due to data availability and consistency requirements, the analysis focuses on the top 30 firms selected from the S&P 500 over a multi-year period. These firms were chosen because they possess relatively complete ESG and financial data across the observation period, allowing reliable panel data analysis and machine learning implementation.

The final dataset consists of 450 observations representing 30 companies observed across 15 years from 2010 to 2024. The dataset contains both ESG indicators and financial performance variables, enabling examination of the relationship between corporate financial strength and ESG performance.

The dataset includes the following variables:

- ESG Score
- Environmental Score (E)
- Social Score (S)
- Governance Score (G)
- Return on Assets (ROA)
- Return on Equity (ROE)
- Total Assets
- Debt
- Company
- Year

The data was imported into RStudio using Excel and CSV file formats. Prior to analysis, the dataset underwent several preprocessing procedures including:

- removal of missing values,
- column name standardization,
- conversion of variables into numeric format,
- and verification of data consistency.

The dataset was structured as balanced panel data for panel regression analysis. Each company contained observations across the same time period, allowing more reliable longitudinal analysis.

The use of panel data improves analytical quality because it captures both cross-sectional and time-series variation within the sample. This allows the study to better examine how financial performance variables influence ESG performance over time.

3.3 Variable Definitions

The study utilizes one dependent variable and multiple independent financial variables.

3.3.1 Dependent Variable

ESG Score

The ESG score represents the overall Environmental, Social, and Governance performance of a company. It reflects corporate sustainability practices, governance quality, environmental responsibility, and social engagement.

In this study, ESG score serves as the dependent variable because the research aims to examine whether financial performance can explain ESG outcomes.

Higher ESG scores indicate stronger sustainability performance and greater commitment toward responsible corporate behavior.

3.3.2 Independent Variables

Return on Assets (ROA)

Return on Assets measures a company's profitability relative to its total assets. It indicates how efficiently a company utilizes its assets to generate earnings.

ROA is calculated as:

$$ROA = \frac{Net\ Income}{Total\ Assets}$$

Higher ROA values generally indicate stronger operational efficiency and profitability.

Return on Equity (ROE)

Return on Equity measures a company's profitability relative to shareholder equity. It reflects management effectiveness in generating returns for investors.

ROE is calculated as:

$$ROE = \frac{Net\ Income}{Shareholder\ Equity}$$

Higher ROE values suggest improved financial performance and stronger shareholder returns.

Total Assets

Total Assets represent the size and resource base of a company. Larger firms often possess greater financial capacity to invest in ESG initiatives, sustainability programs, and governance systems.

Firm size is frequently used within ESG literature because larger corporations tend to face greater public scrutiny and stakeholder pressure regarding sustainability practices.

Debt

Debt represents the financial leverage of a company. Highly leveraged firms may experience financial constraints that influence ESG investment decisions.

Debt levels may positively or negatively influence ESG performance depending on a company's financial strategy and risk management structure.

3.4 Econometric Models

The study applies several econometric techniques in order to examine the relationship between financial performance and ESG performance. Econometric analysis allows statistical evaluation of the significance, direction, and explanatory power of financial variables in predicting ESG outcomes.

The study primarily utilizes:

- Ordinary Least Squares (OLS) regression,
- Multiple Linear Regression,
- and Fixed Effects Panel Regression.

These models were selected because they are widely used within financial and ESG literature for examining relationships between corporate financial indicators and sustainability performance.

3.4.1 Correlation Analysis

Before conducting regression analysis, correlation analysis was performed to examine the strength and direction of the relationship between ESG performance and financial variables.

The Pearson correlation coefficient was used to measure linear association between variables. Correlation values range from:

- -1 indicating a perfect negative relationship,
- 0 indicating no relationship,

- and +1 indicating a perfect positive relationship.

This initial analysis helps identify whether financial indicators such as ROA and ROE exhibit positive or negative associations with ESG performance.

3.4.2 Ordinary Least Squares (OLS) Regression

Ordinary Least Squares (OLS) regression was applied as the primary econometric technique for estimating the relationship between financial performance and ESG performance.

OLS regression estimates the linear relationship between dependent and independent variables by minimizing the sum of squared residuals. The method is widely used in financial research because it provides interpretable coefficient estimates and statistical significance testing.

The simple OLS regression model used in this study is expressed as:

$$ESG_i = \alpha + \beta_1 ROA_i + \varepsilon_i$$

Where:

- ESG represents the dependent variable,
- ROA represents Return on Assets,
- α represents the intercept,
- β represents the regression coefficient,
- and ε represents the error term.

The purpose of this model is to examine whether profitability individually influences ESG performance.

3.4.3 Multiple Linear Regression

A multiple regression model was further applied in order to evaluate the combined effect of multiple financial indicators on ESG performance.

The multiple regression model is specified as:

$$ESG_i = \alpha + \beta_1 ROA_i + \beta_2 ROE_i + \beta_3 Debt_i + \beta_4 Assets_i + \varepsilon_i$$

This model allows simultaneous examination of:

- profitability,
- shareholder return,
- leverage,
- and firm size.

The use of multiple regression improves explanatory power by controlling for several financial characteristics simultaneously.

3.4.4 Fixed Effects Panel Regression

Because the dataset consists of multiple companies observed across several years, panel data regression analysis was also conducted.

The study applies a Fixed Effects panel model using the within estimator. Fixed Effects regression controls for unobserved firm-specific characteristics that remain constant over time, such as managerial culture, industry behaviour, corporate structure, or organizational policies.

The Fixed Effects model is expressed as:

$$ESG_{it} = \alpha_i + \beta X_{it} + \varepsilon_{it}$$

Where:

- i represents individual firms,
- t represents time,
- X_{it} represents financial explanatory variables,
- and α_i captures firm-specific effects.

The Fixed Effects approach was selected because it reduces omitted variable bias arising from unobserved heterogeneity across firms. This improves estimation reliability when analysing longitudinal ESG and financial data.

The panel regression was implemented using the plm package in RStudio with the within transformation method.

3.4.5 Residual Diagnostics

Residual diagnostic analysis was conducted to evaluate the validity and reliability of the regression models.

Several diagnostic plots were generated, including:

- residual versus fitted plots,
- Q-Q plots,
- scale-location plots,
- and residual histograms.

These diagnostics help evaluate:

- normality of residuals,
- homoscedasticity,
- linearity assumptions,
- and potential outliers.

Residual analysis is important for assessing whether the econometric models satisfy the assumptions required for reliable statistical inference.

3.5 Machine Learning Models

In addition to econometric analysis, the study applies multiple machine learning models in order to evaluate predictive performance and identify nonlinear relationships between financial variables and ESG performance.

3.5.1 Random Forest

Random Forest is an ensemble machine learning algorithm based on multiple decision trees. The model combines predictions from numerous trees to improve predictive accuracy and reduce overfitting.

The model is particularly useful for:

- capturing nonlinear relationships,
- identifying variable importance,
- and improving prediction stability.

Random Forest was applied using ROA, ROE, Debt, and Total Assets as predictor variables for ESG performance.

3.5.2 XGBoost

Extreme Gradient Boosting (XGBoost) is an advanced boosting algorithm widely used in predictive analytics and financial forecasting.

The model sequentially improves prediction accuracy by minimizing residual errors across multiple iterations.

XGBoost is highly effective for:

- complex financial datasets,
- nonlinear modelling,
- and high-dimensional prediction tasks.

The study uses XGBoost to evaluate whether advanced machine learning methods outperform traditional regression techniques in ESG prediction.

3.5.3 LASSO Regression

LASSO (Least Absolute Shrinkage and Selection Operator) regression combines regression analysis with automatic variable selection.

The method penalizes weak coefficients and reduces less important variables toward zero.

LASSO was included in the study because it:

- reduces overfitting,
- improves interpretability,
- and identifies the strongest financial predictors of ESG performance.

3.6 Evaluation Metrics

Several evaluation metrics were used to compare model performance across econometric and machine learning approaches.

3.6.1 Root Mean Squared Error (RMSE)

RMSE measures the average magnitude of prediction error between actual and predicted values.

It is calculated as:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

Lower RMSE values indicate stronger predictive accuracy.

3.6.2 Mean Absolute Error (MAE)

MAE measures the average absolute difference between predicted and actual values.

The metric is expressed as:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

Lower MAE values indicate better model performance.

3.6.3 Coefficient of Determination (R^2)

The coefficient of determination measures the proportion of variance explained by the model.

R^2 values range between 0 and 1:

- values closer to 1 indicate stronger explanatory power,
- while lower values indicate weaker model fit.

The study compares R^2 values across Linear Regression, Random Forest, XGBoost, and LASSO models to determine the strongest predictive framework.

3.6.4 Variable Importance Analysis

Variable importance analysis was conducted within Random Forest and XGBoost models in order to identify the most influential predictors of ESG performance.

This analysis helps determine whether profitability, leverage, or firm size contributes most significantly to ESG prediction.

3.6.5 Model Comparison

Finally, all models were compared using RMSE, MAE, and R^2 metrics in order to identify the best-performing analytical framework.

The comparison allows evaluation of whether machine learning approaches outperform traditional econometric methods in explaining ESG performance.

Chapter 4. Empirical Results & Discussion

4.1 Introduction

This chapter presents the empirical findings of the study and discusses the relationship between ESG performance and corporate financial indicators through multiple analytical stages. The empirical analysis was conducted progressively in order to examine the ESG–financial relationship from different perspectives and to improve the robustness of the research framework.

The study initially began with a traditional approach commonly found within existing ESG literature, where ESG performance was examined as an explanatory variable affecting corporate financial performance. In the first stage, a short-term one-year observation period was analysed in order to evaluate whether ESG scores significantly influenced financial indicators such as Return on Assets (ROA) and Return on Equity (ROE). This preliminary analysis provided initial insight into the relationship between sustainability performance and financial outcomes.

Following the initial findings, the analysis was extended into a multi-year panel framework in order to improve longitudinal reliability and capture variations across firms and time periods. The second stage therefore examined the impact of ESG performance on financial indicators using panel data observations across multiple years for selected companies from the S&P 500 index. The use of panel data allowed a more comprehensive examination of firm-specific and time-related effects within the ESG–financial relationship.

However, despite extending the observation period and applying panel econometric techniques, the results suggested that the impact of ESG performance on financial outcomes remained relatively limited and inconsistent across certain models. These findings motivated a reconsideration of the conventional ESG research direction and led to the development of the reversed analytical framework adopted in the final stage of the study.

The third and most important stage of the empirical analysis reversed the traditional relationship by examining financial performance as the explanatory factor and ESG

performance as the dependent variable. This reversed framework was developed in order to investigate whether financially stronger firms possess greater capability to invest in sustainability practices, governance systems, and ESG-related initiatives.

To evaluate this reversed relationship, the study applied both econometric and machine learning techniques. Traditional econometric methods included Ordinary Least Squares (OLS) regression, multiple regression analysis, and Fixed Effects panel regression. In addition, several machine learning models, including Random Forest, XGBoost, and LASSO regression, were implemented in order to compare predictive performance and identify the strongest determinants of ESG performance.

The chapter therefore presents not only the statistical findings of the models but also a broader discussion regarding the evolution of the research framework, the consistency of empirical evidence, and the comparative performance of econometric and machine learning approaches. Furthermore, all figures, plots, and model outputs are individually interpreted in order to provide detailed analytical discussion and stronger connection between theoretical expectations and empirical observations.

Overall, the empirical chapter aims to demonstrate how the study evolved from a conventional ESG-financial framework toward a more comprehensive and predictive analysis of the financial determinants of ESG performance.

4.2 Initial ESG → Financial Performance Analysis (One-Year Observation)

4.2.1 Variable Structure and Model Design

The first stage of the empirical analysis focused on examining the traditional relationship between ESG performance and corporate financial performance using a limited one-year observation period. In this initial framework, ESG performance was treated as the independent variable, while financial performance indicators such as Return on Assets (ROA) served as dependent variables.

The purpose of this preliminary analysis was to evaluate whether ESG performance had an observable short-term influence on corporate profitability among the selected firms from the S&P 500 sample.

Table 4.1

Variable	Type	Description	Role in Initial Model
ESG Score	Independent Variable	Composite Environmental, Social, and Governance performance indicator measuring overall corporate sustainability performance	Main Explanatory Variable
Environmental Score (E)	Independent Variable	Measures firm environmental sustainability practices and environmental risk management	ESG Pillar Variable
Social Score (S)	Independent Variable	Measures social responsibility, labor practices, and stakeholder engagement	ESG Pillar Variable
Governance Score (G)	Independent Variable	Measures governance quality, transparency, and board effectiveness	ESG Pillar Variable
ROA (Return on Assets)	Dependent Variable	Measures efficiency of asset utilization in generating profits	Financial Performance Indicator
ROE (Return on Equity)	Dependent Variable	Measures profitability generated from shareholders' equity	Financial Performance Indicator
Total Assets	Control Variable	Proxy for firm size and operational scale	Firm Size Control
Debt	Control Variable	Represents firm leverage and debt obligations	Financial Risk Control
Company	Cross-sectional Identifier	Identifies firms included in the sample	Firm Identifier

Table 4.1 presents the variables used in the initial empirical model, including their classification, descriptions, and analytical roles within the regression framework.

The ESG score was used as the primary explanatory variable because it represents the overall sustainability performance of firms through environmental, social, and governance dimensions. In addition, ESG pillar variables including Environmental Score (E), Social Score (S), and Governance Score (G) were incorporated to allow more detailed sustainability analysis.

Financial indicators such as ROA and ROE were included to evaluate profitability performance, while Total Assets and Debt were used as control variables representing

firm size and leverage characteristics respectively. Company identifiers were also included to distinguish firms within the dataset.

The variable structure demonstrates that the initial model was primarily designed to test whether stronger ESG performance contributes to improved short-term financial outcomes.

4.2.2 Distribution of ESG Scores

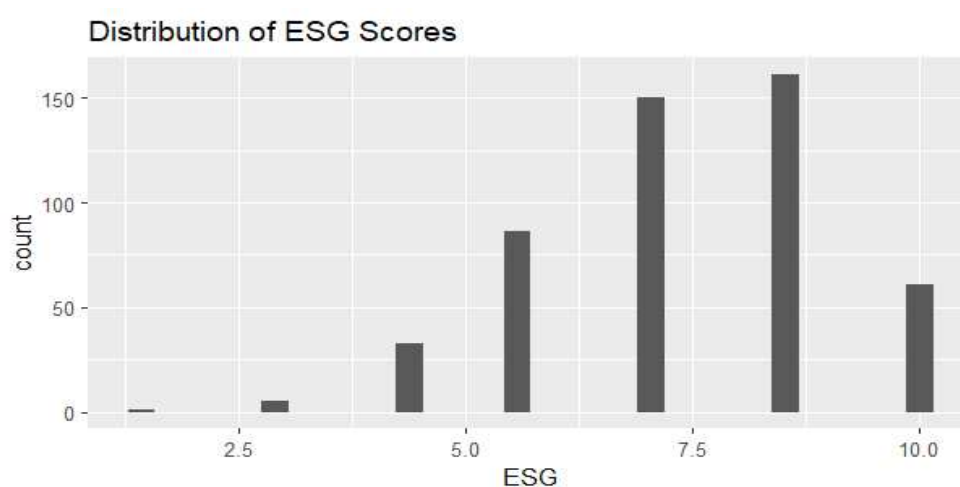


Figure 4.1

Figure 4.1 illustrates the distribution of ESG scores across the sample firms included within the one-year observation dataset. The histogram demonstrates that the majority of companies possess ESG scores concentrated within the medium-to-high range, particularly between scores of 6 and 9.

This distribution suggests that most firms within the selected S&P 500 sample already maintain relatively developed sustainability and governance practices. Very low ESG scores appear infrequently within the dataset, indicating that large publicly listed firms generally exhibit moderate to strong ESG engagement.

The concentration of observations around higher ESG values also reflects the increasing importance of sustainability disclosure and ESG integration within large corporations. Since the sample consists of major S&P 500 firms, the relatively elevated ESG performance is consistent with expectations regarding large-cap

companies facing stronger stakeholder pressure, regulatory attention, and investor scrutiny related to sustainability performance.

Overall, the ESG distribution indicates relatively limited variation between firms, which may partially explain the weaker statistical relationships observed within the initial short-term regression analysis.

4.2.3 Distribution of Financial Performance (ROA)

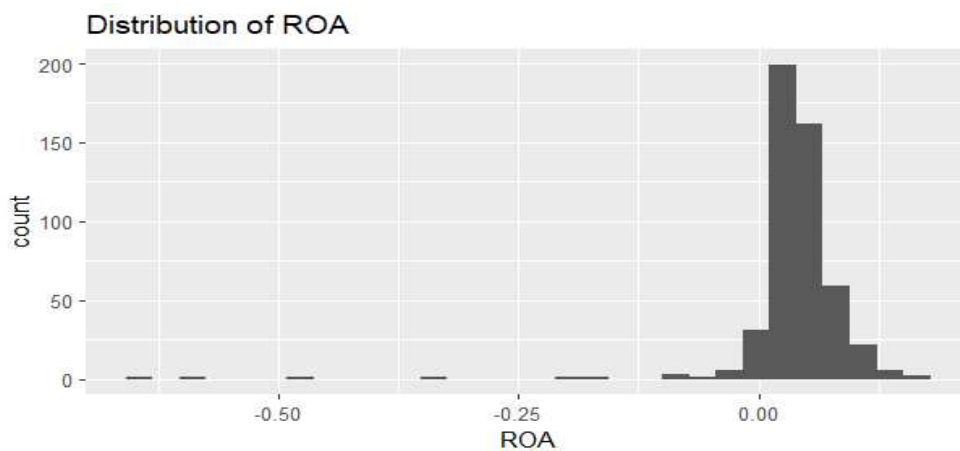


Figure 4.2

Figure 4.2 presents the distribution of Return on Assets (ROA) within the sample firms. The histogram demonstrates that the majority of observations are concentrated around small positive profitability values, while several firms exhibit negative ROA values representing financial losses or weaker operational efficiency.

The distribution appears moderately skewed, with a number of outlier observations located within the negative profitability range. This indicates that although many firms maintained stable profitability during the observation period, financial performance varied considerably across companies.

The existence of both profitable and less profitable firms improves the analytical quality of the dataset because it allows examination of ESG relationships across different financial conditions. However, the presence of outliers and profitability dispersion may

also reduce the strength of simple linear relationships between ESG performance and short-term profitability indicators.

The ROA distribution therefore suggests that financial performance during the observed period was relatively heterogeneous across firms, despite similarities in ESG performance levels.

4.2.4 Correlation Analysis

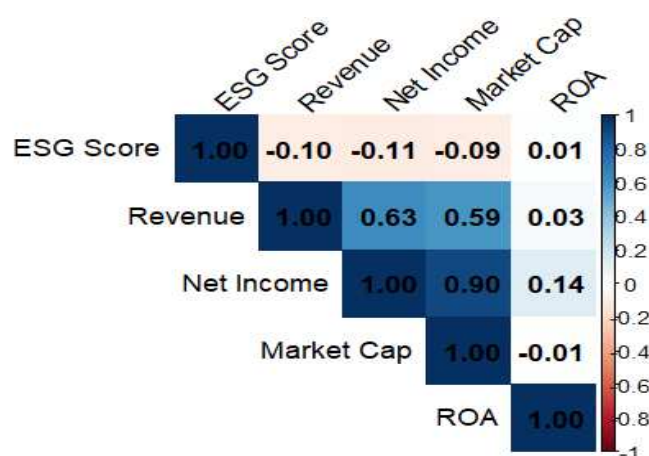


Figure 4.3

Figure 4.3 presents the correlation matrix heatmap illustrating the relationships between ESG Score, Revenue, Net Income, Market Capitalization, and Return on Assets (ROA) within the one-year observation framework.

The results demonstrate that ESG Score exhibited extremely weak relationships with the selected financial variables. In particular, the correlation between ESG Score and ROA was approximately 0.01, indicating an almost non-existent linear relationship between sustainability performance and profitability during the observed period. Similarly, ESG Score showed weak negative correlations with Revenue (-0.10), Net Income (-0.11), and Market Capitalization (-0.09).

These findings suggest that short-term ESG performance did not exhibit a meaningful direct association with immediate financial outcomes among the sampled S&P 500 firms. The weak correlations support the earlier regression results indicating limited explanatory power of ESG variables within the one-year framework.

In contrast, stronger relationships were observed among the financial variables themselves. Revenue demonstrated a relatively strong positive correlation with Net Income (0.63) and Market Capitalization (0.59), while Net Income and Market Capitalization exhibited a particularly high positive correlation of approximately 0.90. These results are economically intuitive because larger firms with stronger revenues generally tend to report higher profitability and market valuation.

The correlation matrix therefore reveals an important distinction within the dataset. While financial indicators appear strongly interconnected, ESG performance does not demonstrate a substantial short-term relationship with profitability or market-related variables within the selected observation period.

From an analytical perspective, these findings further strengthened the assumption that ESG-related financial effects may require longer time horizons to become observable. The results also provided an early indication that the traditional ESG → financial performance framework may possess limited explanatory capacity within short-term empirical settings.

Consequently, the weak correlation structure observed in the initial analysis contributed to the decision to extend the study toward multi-year panel analysis in the following stage of the research.

4.2.5 Simple Regression Results

```

--- SIMPLE REGRESSION ---
> summary(model1)

Call:
lm(formula = ROA ~ ESG, data = data)

Residuals:
    Min       1Q   Median       3Q      Max
-0.68399 -0.01237  0.00067  0.02031  0.12578

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.0371690   0.0128259   2.898  0.00392 **
ESG           0.0002357   0.0016783   0.140  0.88837
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.06154 on 495 degrees of freedom
Multiple R-squared:  3.984e-05, Adjusted R-squared: -0.00198
F-statistic: 0.01972 on 1 and 495 DF, p-value: 0.8884

```

Image 4.1

The simple regression model was estimated to examine whether ESG performance

significantly influenced Return on Assets (ROA) during the one-year observation period:

$$ROA = \beta_0 + \beta_1 ESG + \epsilon$$

Key Findings

- The ESG coefficient was positive but extremely small, indicating only a negligible positive relationship between ESG performance and ROA.
- The p-value of the ESG coefficient ($p = 0.888$) was statistically insignificant, suggesting that ESG performance did not meaningfully explain variations in profitability within the sample firms.
- The R^2 value was approximately 0.00004, indicating that ESG performance explained virtually none of the variation in ROA.
- The regression model therefore demonstrated extremely weak explanatory power within the short-term framework.
- Residual values remained relatively dispersed, further indicating weak predictive capability of the model.

Interpretation

- The findings suggest that ESG performance alone may not generate immediate short-term financial benefits measurable through ROA.
- Sustainability investments often require longer implementation periods before operational or financial improvements become observable.
- ESG initiatives may initially involve additional costs related to environmental management, governance restructuring, reporting systems, and social responsibility activities.
- Consequently, the short-term financial effect of ESG performance may remain limited or delayed.

Theoretical Implication

- The results support portions of the existing literature arguing that ESG-financial relationships are frequently long-term rather than immediate.

- The findings also indicate that short observational periods may be insufficient for capturing the broader economic effects of sustainability practices.
- This initial evidence motivated the extension of the analysis toward longer panel observations in the subsequent stage of the research.

4.2.6 Multiple Regression Results

```

--- MULTIPLE REGRESSION ---
> summary(model2)

Call:
lm(formula = ROA ~ ESG + data[["Market Cap"]], data = data)

Residuals:
    Min       1Q   Median       3Q      Max
-0.68413 -0.01211  0.00065  0.02056  0.12567

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.749e-02  1.301e-02   2.882  0.00412 **
ESG           2.126e-04  1.687e-03   0.126  0.89977
data[["Market Cap"]] -1.379e-15  9.061e-15 -0.152  0.87908
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.0616 on 494 degrees of freedom
Multiple R-squared:  8.674e-05, Adjusted R-squared:  -0.003961
F-statistic: 0.02143 on 2 and 494 DF,  p-value: 0.9788

```

Image 4.2

The multiple regression model extended the initial analysis by incorporating Market Capitalization alongside ESG performance in order to evaluate whether firm size improved explanatory power for ROA.

$$ROA = \beta_0 + \beta_1 ESG + \beta_2 Size + \epsilon$$

Key Findings

- ESG remains **statistically insignificant** ($p \approx 0.90$)
- The model's explanatory power remains extremely low ($R^2 \approx 0.00009$)
- Both ESG Score and Market Capitalization exhibited statistically insignificant coefficients.
- The ESG coefficient remained extremely small and insignificant even after controlling for Market Capitalization.

- Market Capitalization also failed to demonstrate a meaningful relationship with ROA during the observed period.
- The model generated an extremely low R^2 value close to zero, indicating very weak overall explanatory performance.
- The F-statistic was statistically insignificant, suggesting that the regression model as a whole lacked predictive strength.

Interpretation

- The inclusion of Market Capitalization did not materially improve the relationship between ESG performance and profitability.
- These findings suggest that neither sustainability performance nor firm size alone sufficiently explained profitability differences within the one-year framework.
- The results may indicate that short-term profitability is influenced more strongly by operational, industry-specific, macroeconomic, or managerial factors rather than ESG performance alone.
- The weak model fit also suggests that the ESG-financial relationship may involve more complex and nonlinear dynamics that cannot be fully captured through a limited short-term regression framework.

Analytical Implication

- The persistence of weak statistical relationships after introducing control variables reinforced the limitations of the short-term ESG → financial performance framework.
- The findings provided empirical justification for expanding the research into a multi-year panel structure capable of capturing time dynamics and firm-level heterogeneity.
- This stage therefore served as an important transition point within the study, ultimately contributing to the later reversal of the analytical framework toward examining financial performance as a determinant of ESG outcomes.

4.3 Extended ESG → Financial Performance Panel Analysis

4.3.1 Introduction to the Extended Panel Analysis

Following the initial one-year analysis, the study was extended to a broader longitudinal framework in order to obtain a more comprehensive understanding of the relationship between ESG performance and financial outcomes. While the first-stage analysis provided a short-term snapshot of the association between ESG scores and profitability indicators, it was limited in its ability to capture changes occurring across time. Financial performance and sustainability behaviour are dynamic in nature, and therefore a longer observation period is necessary to examine whether ESG practices produce stable and measurable financial implications over multiple years.

To address this limitation, the present section adopts a ten-year panel data framework based on 30 selected firms from the S&P 500 index. The panel structure combines both cross-sectional and time-series dimensions, allowing the analysis to evaluate firm behaviour across different years while controlling for firm-specific characteristics. Compared to a single-period analysis, panel data provides stronger statistical reliability and improves the ability to identify persistent trends within ESG and financial performance relationships.

The extended analysis continues to examine ESG score as the principal explanatory variable, while Return on Assets (ROA) and Return on Equity (ROE) are considered the primary indicators of financial performance. In addition, ESG pillar scores — Environmental (E), Social (S), and Governance (G) — are analysed separately in order to determine whether specific sustainability dimensions influence firm profitability differently. Control variables such as Total Assets and Debt are also incorporated to account for differences in firm size and leverage.

This section employs multiple analytical approaches, including correlation analysis, Ordinary Least Squares (OLS) regression, ESG pillar regression, and fixed-effects panel regression models. The use of fixed-effects estimation is particularly important because it controls for unobserved firm-level heterogeneity that may influence financial outcomes over time. Diagnostic procedures and graphical analysis are also included to evaluate model validity and data behaviour. The objective of this section is not only to determine whether ESG influences financial performance over a longer period, but also to evaluate the stability, strength, and consistency of this relationship within a panel

data environment. The findings obtained from this stage later serve as the foundation for the reversed analytical framework presented in the subsequent section of the thesis, where financial variables are examined as determinants of ESG performance.

4.3.2 Variable Structure and Dataset Framework

Variable	Type	Description	Role in Revised Model
ESG Score	Dependent Variable	Composite Environmental, Social, and Governance score representing overall sustainability performance	Target Variable
ROA (Return on Assets)	Independent Variable	Measures profitability generated from firm assets	Profitability Predictor
ROE (Return on Equity)	Independent Variable	Measures return generated on shareholders' equity	Profitability Predictor
Debt	Independent Variable	Represents leverage and financial obligations of the firm	Financial Structure Predictor
Total Assets	Independent Variable	Proxy for firm size and operational capacity	Firm Size Predictor
Environmental Score (E)	Supplementary ESG Variable	Environmental sustainability performance indicator	ESG Subcomponent
Social Score (S)	Supplementary ESG Variable	Social responsibility and stakeholder management indicator	ESG Subcomponent
Governance Score (G)	Supplementary ESG Variable	Corporate governance quality indicator	ESG Subcomponent
Company	Panel Identifier	Identifies firms across multiple observations	Cross-sectional Dimension
Year	Time Identifier	Represents yearly observations across the study period	Time Dimension

Table 4.2

Table 4.3 presents the structure of variables used within the ten-year panel analysis. In this stage, ESG score remained the primary explanatory variable, while ROA and ROE served as indicators of firm profitability. Debt and Total Assets were included as control variables to account for leverage and firm-size effects.

The Environmental (E), Social (S), and Governance (G) scores were further introduced as supplementary ESG dimensions in order to evaluate the individual contribution of each sustainability component toward financial performance. The inclusion of Company and Year identifiers enabled the construction of a balanced panel dataset, thereby allowing the implementation of fixed-effects panel regression models.

Compared to the initial one-year framework, the extended dataset provides a more robust and comprehensive structure for evaluating the long-term interaction between sustainability performance and corporate profitability.

4.3.3 Correlation Matrix Analysis

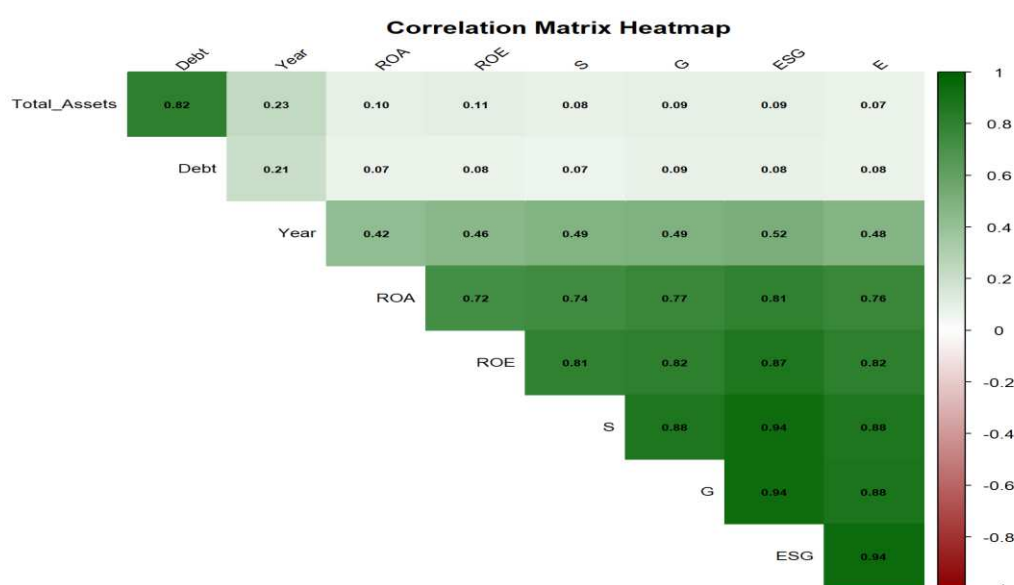


Figure 4.4

The correlation heatmap illustrates the pairwise relationships among ESG indicators, profitability variables, control variables, and ESG pillar scores. The findings reveal considerably stronger associations than those observed in the first-stage analysis.

The correlation between ESG and ROA was approximately 0.81, while the relationship between ESG and ROE reached approximately 0.87. These values indicate a strong positive relationship between sustainability performance and firm profitability across the observation period. In contrast to the earlier short-term analysis, where the ESG–financial relationship appeared weak and statistically insignificant, the longitudinal panel framework demonstrates a substantially stronger association.

The Environmental, Social, and Governance pillar scores also exhibited very high positive correlations with the aggregate ESG score, with coefficients exceeding 0.88 in several cases. This confirms the internal consistency of the ESG measurement framework and suggests that all sustainability dimensions contribute meaningfully toward overall ESG performance.

Among the control variables, Debt and Total Assets displayed relatively weaker direct relationships with ESG indicators, although Total Assets and Debt themselves were strongly correlated. This suggests that larger firms generally maintain higher leverage levels but that leverage alone does not strongly determine ESG performance.

Overall, the correlation matrix provides preliminary evidence that ESG performance is positively associated with long-term profitability among the selected S&P 500 firms.

4.3.4 Scatter Plot Analysis

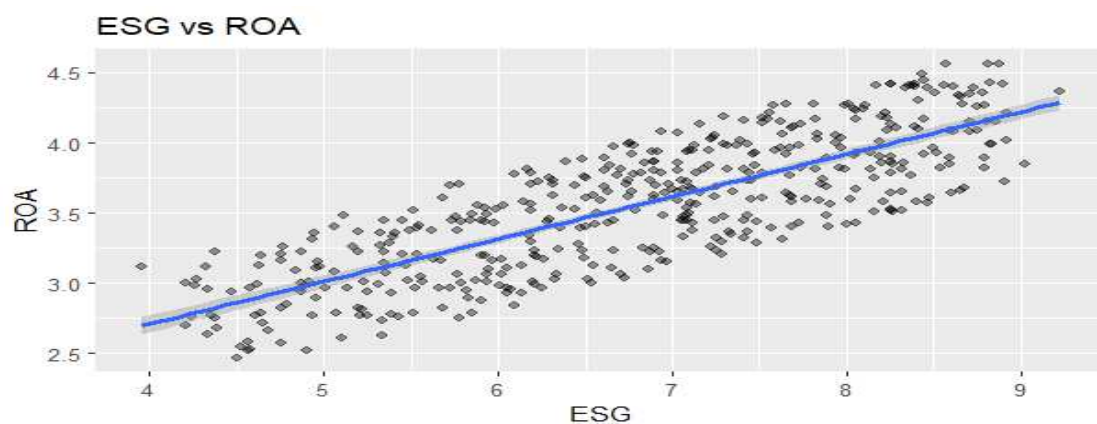


Figure 4.5

The scatter plot examining ESG and ROA demonstrates a clear positive linear relationship between sustainability performance and operational profitability. The upward-sloping

regression line indicates that firms with higher ESG scores generally report higher returns on assets.

Compared to the first-stage analysis, the observations in the extended dataset appear more concentrated around the fitted regression line, suggesting improved model consistency and stronger explanatory behaviour over the longer observation period. Although some dispersion remains visible, the general trend strongly supports the existence of a positive ESG–ROA relationship.

The graphical evidence aligns closely with the correlation and regression findings and visually confirms the strengthening of the ESG–profitability relationship within the panel framework.

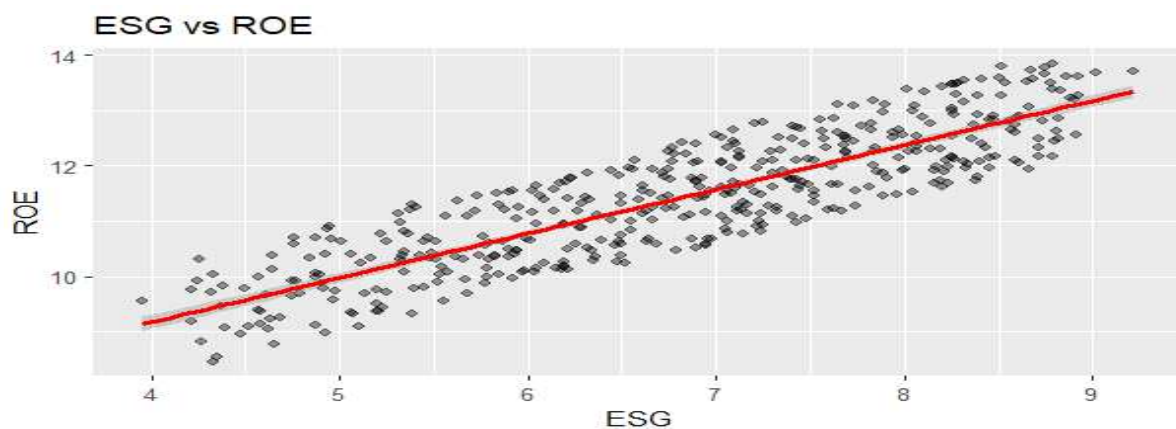


Figure 4.6

The relationship between ESG and Return on Equity (ROE) appears even stronger than the ESG–ROA relationship. The scatter plot shows a highly consistent upward trend, indicating that firms with stronger ESG performance tend to generate superior returns for shareholders.

The fitted regression line demonstrates a stable positive slope with comparatively lower dispersion around the trend line. This suggests that ESG performance may have a stronger association with shareholder profitability than with operational profitability.

The graphical findings therefore reinforce the argument that sustainability performance contributes positively toward firm value creation over time.

4.3.5 Ordinary Least Squares (OLS) Regression Analysis

```
===== ESG → ROA =====
> model_roa <- lm(ROA ~ ESG, data = data)
> summary(model_roa)

Call:
lm(formula = ROA ~ ESG, data = data)

Residuals:
    Min       1Q   Median       3Q      Max
-0.50698 -0.22320  0.01629  0.22879  0.48260

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.50230    0.07188   20.9   <2e-16 ***
ESG           0.30207    0.01038   29.1   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2766 on 448 degrees of freedom
Multiple R-squared:  0.654,    Adjusted R-squared:  0.6533
F-statistic: 846.9 on 1 and 448 DF, p-value: < 2.2e-16
```

Image 4.3 ESG → ROA Regression Output

The first regression model examined the direct impact of ESG performance on Return on Assets (ROA). The regression coefficient for ESG was positive and highly statistically significant ($p < 0.001$), indicating that improvements in ESG performance are associated with higher operational profitability.

The model produced an R^2 value of approximately 0.65, suggesting that ESG performance explains a substantial proportion of the variation in ROA across the selected firms. Compared to the initial short-term analysis, where the ESG coefficient was statistically insignificant, the extended panel analysis demonstrates a much stronger and more stable relationship.

The findings indicate that sustainability performance may contribute positively toward operational efficiency and asset utilization over the long term.

```
===== ESG → ROE =====
> model_roe <- lm(ROE ~ ESG, data = data)
> summary(model_roe)

Call:
lm(formula = ROE ~ ESG, data = data)

Residuals:
    Min       1Q   Median       3Q      Max
-0.97908 -0.49101 -0.01086  0.47952  1.02737

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  5.97576    0.14830   40.29   <2e-16 ***
ESG           0.79984    0.02142   37.35   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5708 on 448 degrees of freedom
Multiple R-squared:  0.7569,    Adjusted R-squared:  0.7564
F-statistic: 1395 on 1 and 448 DF, p-value: < 2.2e-16
```

Image 4.4 ESG → ROE Regression Output

The second regression model evaluated the relationship between ESG performance and Return on Equity (ROE). Similar to the ROA model, the ESG coefficient remained positive and highly significant.

The model generated an R^2 value exceeding 0.75, indicating that ESG explains a large proportion of the variation in shareholder profitability. The statistical significance and explanatory power of the model suggest that firms with stronger ESG practices tend to deliver higher returns to equity investors.

The ROE model therefore provides additional support for the argument that sustainability performance contributes positively toward financial outcomes.

4.3.6 ESG Pillar Analysis

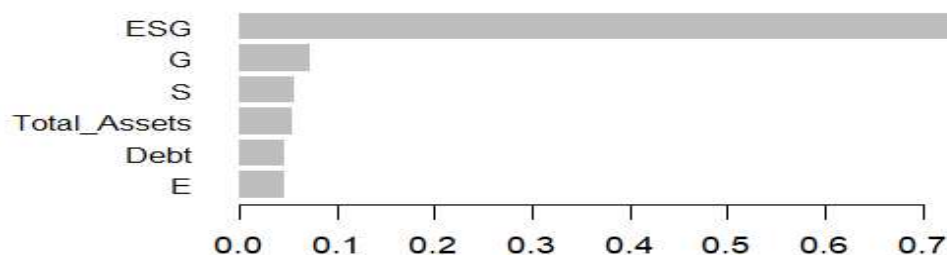


Figure 4.7 ESG Pillar Importance Visualization

The ESG pillar importance chart demonstrates the relative contribution of the Environmental (E), Social (S), and Governance (G) dimensions toward overall model performance. The aggregate ESG score emerged as the strongest contributor, while Governance and Social dimensions also displayed meaningful explanatory power.

The figure suggests that sustainability performance is multidimensional in nature and that different ESG components contribute differently toward financial outcomes.

```

===== ESG Pillars → ROA =====
> model_pillars <- lm(ROA ~ E + S + G, data = data)
> summary(model_pillars)

Call:
lm(formula = ROA ~ E + S + G, data = data)

Residuals:
    Min       1Q   Median       3Q      Max
-0.68199 -0.23981  0.02169  0.24595  0.60628

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.61639    0.07282   22.198  < 2e-16 ***
E             0.10420    0.02396    4.348  1.70e-05 ***
S             0.04711    0.02355    2.001   0.046  *
G             0.13332    0.02424    5.501  6.37e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2886 on 446 degrees of freedom
Multiple R-squared:  0.6251,    Adjusted R-squared:  0.6225
F-statistic: 247.8 on 3 and 446 DF,  p-value: < 2.2e-16

```

Image 4.5 ESG Pillars → ROA Regression

The ESG pillar regression for ROA indicates that all three sustainability dimensions positively influence operational profitability. Governance displayed the strongest coefficient among the three pillars, followed by Environmental performance, while the social pillar exhibited a smaller but still statistically significant contribution.

The results imply that firms with stronger governance structures and environmental practices are more likely to achieve higher operational efficiency and profitability levels.

```

===== ESG Pillars → ROE =====
> pillar_model <- lm(ROE ~ E + S + G, data = data)
> summary(pillar_model)

Call:
lm(formula = ROE ~ E + S + G, data = data)

Residuals:
    Min       1Q   Median       3Q      Max
-1.39398 -0.47722 -0.03138  0.49902  1.37823

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  6.28095    0.15508   40.502  < 2e-16 ***
E             0.25663    0.05104    5.028  7.18e-07 ***
S             0.23090    0.05014    4.605  5.39e-06 ***
G             0.26624    0.05161    5.158  3.76e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.6147 on 446 degrees of freedom
Multiple R-squared:  0.7193,    Adjusted R-squared:  0.7174
F-statistic: 381 on 3 and 446 DF,  p-value: < 2.2e-16

```

Image 4.6 ESG Pillars → ROE Regression

The ESG pillar regression for ROE similarly revealed strong positive and statistically significant relationships across all three ESG dimensions. Governance again emerged as one of the strongest contributors to shareholder profitability.

The consistency of the ESG pillar findings across both ROA and ROE models strengthens the reliability of the overall ESG–financial performance relationship identified within the study.

4.3.7 Panel Regression Analysis

```
===== PANEL MODEL =====
> panel_data <- pdata.frame(data, index = c("Company", "Year"))
> panel_model <- plm(ROA ~ ESG, data = panel_data, model = "within")
> summary(panel_model)
Oneway (individual) effect within Model

Call:
plm(formula = ROA ~ ESG, data = panel_data, model = "within")

Balanced Panel: n = 30, T = 15, N = 450

Residuals:
    Min.    1st Qu.    Median     3rd Qu.     Max.
-0.540301 -0.217819  0.010206  0.216048  0.591941

Coefficients:
      Estimate Std. Error t-value Pr(>|t|)
ESG 0.309725    0.019656  15.757 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares:    51.948
Residual Sum of Squares: 32.619
R-Squared:                0.37208
Adj. R-Squared: 0.32712
F-statistic: 248.284 on 1 and 419 DF, p-value: < 2.22e-16
```

Image 4.7 Fixed Effects Panel Regression Output

The fixed-effects panel regression model was estimated to account for unobservable firm-specific characteristics across the ten-year observation period. The results indicate that ESG performance maintained a positive and highly significant effect on ROA even after controlling for firm-level heterogeneity.

The balanced panel consisted of 30 firms observed across 15 years, producing a total of 450 observations. The ESG coefficient remained statistically significant at the 1% level, confirming that the ESG–profitability relationship persists within firms over time.

Although the R^2 value in the panel framework was lower than in the simple OLS model, this outcome is expected within fixed-effects estimation due to the stricter within-firm variation structure. Nevertheless, the model still provides strong evidence supporting the positive long-term effect of ESG performance on financial outcomes.

4.3.8 Time Effect Analysis

```
==== TIME EFFECT =====
> trend_model <- lm(ROA ~ ESG + Year + ESG:Year, data = data)
> summary(trend_model)

Call:
lm(formula = ROA ~ ESG + Year + ESG:Year, data = data)

Residuals:
    Min       1Q   Median       3Q      Max
-0.50694 -0.22915  0.01646  0.23285  0.48351

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.103e+01  3.504e+01  -0.315   0.753
ESG           1.827e+00  5.042e+00   0.362   0.717
Year          6.223e-03  1.738e-02   0.358   0.720
ESG:Year     -7.569e-04  2.500e-03  -0.303   0.762

Residual standard error: 0.2772 on 446 degrees of freedom
Multiple R-squared:  0.6542,    Adjusted R-squared:  0.6518
F-statistic: 281.2 on 3 and 446 DF,  p-value: < 2.2e-16
```

Image 4.8 ESG × Year Interaction Model

The time-effect regression model was estimated to determine whether the relationship between ESG and ROA changed across the observation period. The model incorporated ESG, Year, and an ESG×Year interaction term.

The results indicate that neither the Year variable nor the interaction term was statistically significant. This suggests that the positive relationship between ESG and profitability remained relatively stable throughout the study period rather than strengthening or weakening over time.

The findings therefore imply that ESG performance exerted a consistently positive influence on profitability across the entire observation horizon.

4.3.9 Residual Diagnostics and Model Validation

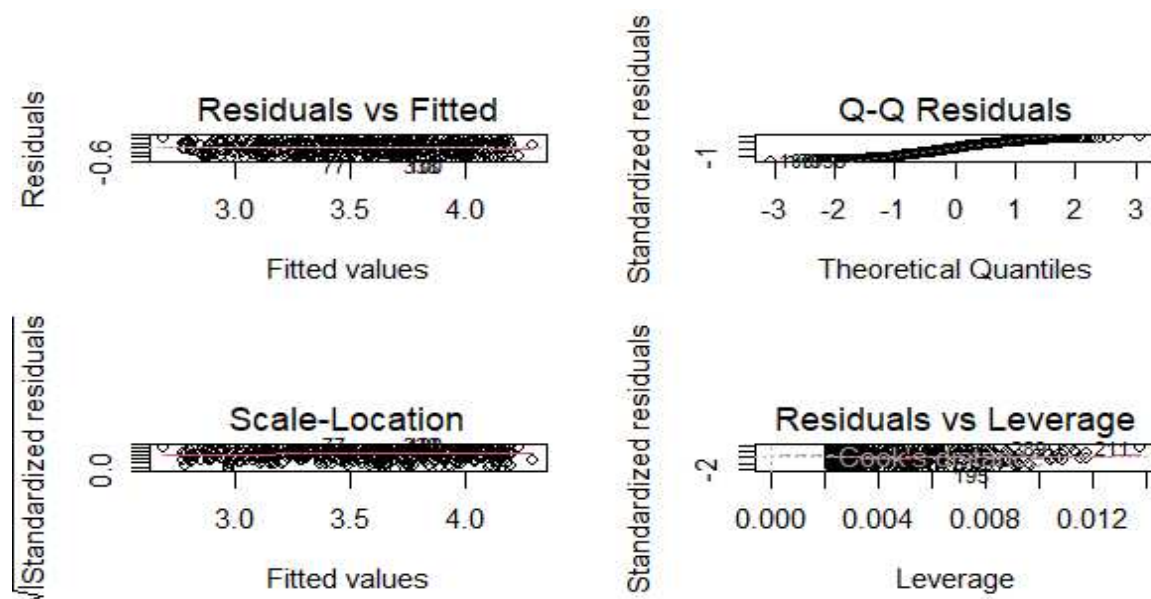


Figure 4.7 Residual Diagnostic Plots

Residual diagnostic testing was conducted to evaluate the validity and reliability of the regression models. The residual-versus-fitted plots did not reveal major nonlinear patterns, suggesting that the linear specification remained generally appropriate. Residual analysis is a critical component of regression diagnostics, used to evaluate whether the assumptions of the linear regression model are satisfied. Residuals represent the difference between observed values and predicted values:

$$\text{Residual} = y_{\text{observed}} - y_{\text{predicted}}$$

If a pattern (e.g., curve) were present, it would suggest that the relationship between ESG and financial performance is non-linear. However, the current plot supports the use of a linear model

The Q-Q plot indicated approximate normality of residuals, although slight deviations from the reference line were visible at the distribution tails. The scale-location plot also suggested acceptable homoscedasticity behaviour across fitted values.

Overall, the diagnostic analysis supports the adequacy of the econometric models used within the study and indicates that the regression assumptions were reasonably satisfied.

4.3.10 Random Forest Analysis

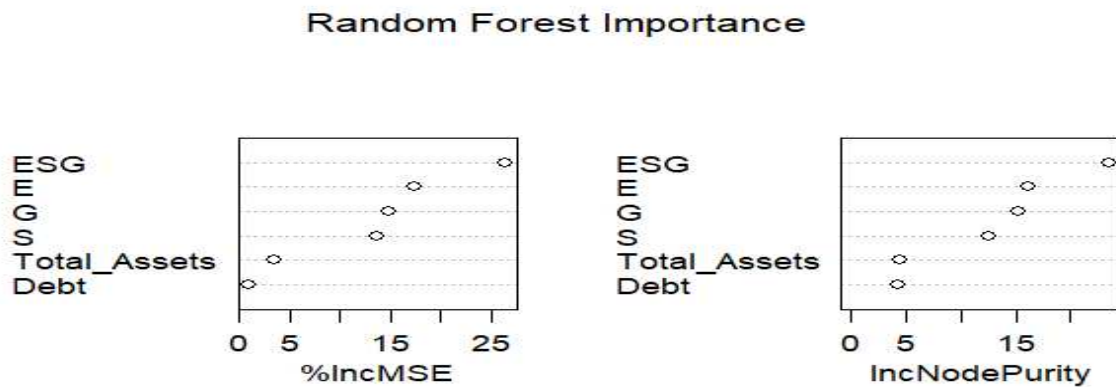


Figure 4.8 Random Forest Variable Importance

The Random Forest variable importance plot identified ESG as the most influential predictor within the machine learning framework. Governance, Social, and Environmental pillar scores also demonstrated meaningful predictive importance, while Debt and Total Assets contributed comparatively less explanatory value.

The Random Forest results therefore reinforce the earlier regression findings by confirming the strong predictive relevance of sustainability indicators for financial performance.

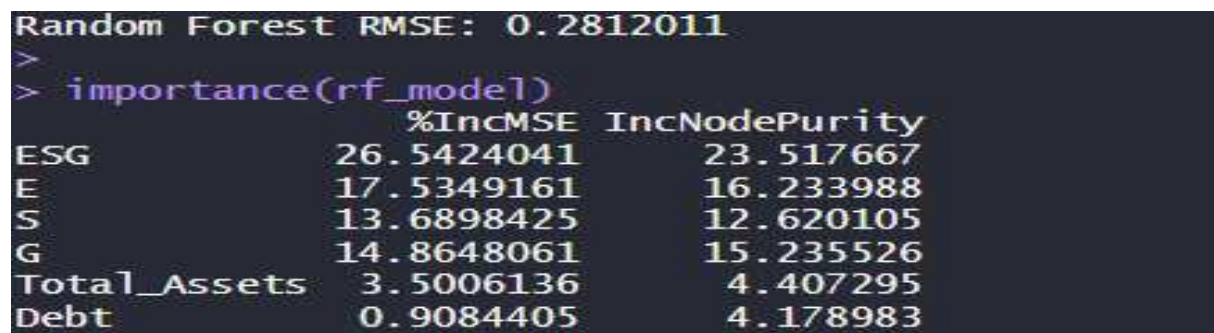


Image 4.9: Random Forest RMSE and Predictor Importance

The Random Forest model generated a relatively low RMSE value, indicating strong predictive performance within the dataset. ESG emerged as the dominant explanatory variable, followed by the Environmental and Governance dimensions.

The results suggest that machine learning approaches are capable of capturing complex nonlinear relationships between ESG variables and profitability indicators.

4.3.11 XGBoost and LASSO Analysis

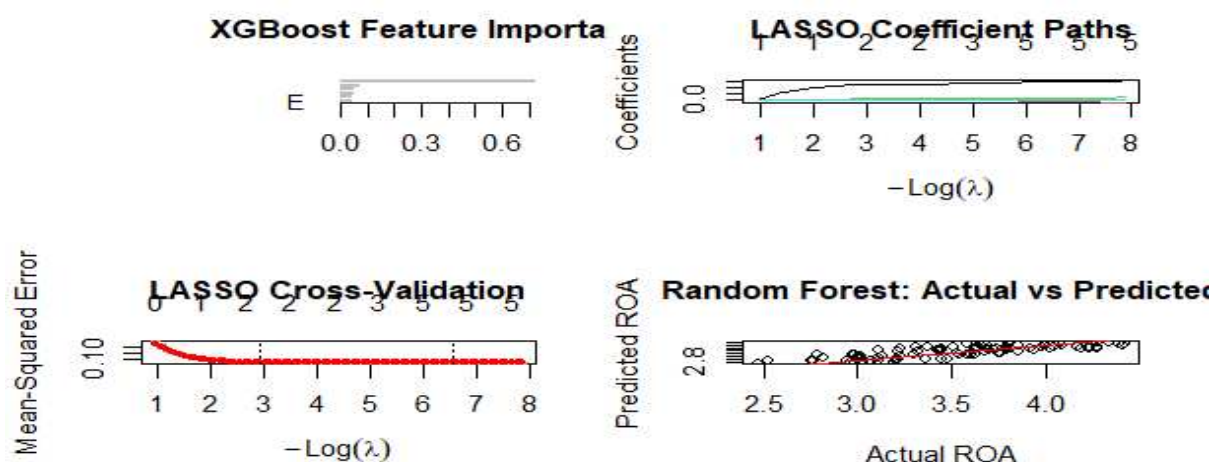


Figure 4.9 XGBoost and LASSO Outputs

The XGBoost model further confirmed the predictive significance of ESG-related variables within the profitability framework. Feature-importance analysis demonstrated that ESG-related dimensions remained dominant contributors within the prediction structure.

The LASSO regression coefficient paths illustrated how weaker predictors were gradually penalized and removed from the model as regularization increased. ESG-related variables consistently retained non-zero coefficients, indicating strong explanatory stability.

The LASSO cross-validation plot further demonstrated the robustness of the selected regularization parameter and confirmed the reliability of the model selection process.

The machine learning findings overall complement the econometric analysis and provide additional evidence supporting the importance of ESG performance in explaining long-term corporate profitability.

4.3.12 Comparative Discussion of Findings

The ten-year panel analysis produced substantially stronger results than the earlier one-year framework. While the short-term analysis indicated weak and statistically insignificant ESG–financial relationships, the extended longitudinal analysis demonstrated strong positive and

highly significant associations across correlation analysis, OLS regression, ESG pillar models, panel regression, and machine learning approaches.

The findings suggest that ESG effects may become financially observable only over longer periods, as sustainability investments often generate gradual strategic, operational, and reputational benefits rather than immediate short-term gains.

Overall, the extended panel analysis strongly supports the argument that ESG performance contributes positively toward corporate profitability, operational efficiency, and shareholder value creation over time.

4.4 Financial Performance as a Determinant of ESG Performance

4.4.1 Introduction to the Reversed Framework

The earlier sections of this study primarily examined the traditional relationship between Environmental, Social, and Governance (ESG) performance and corporate financial performance. Most existing literature in sustainable finance focuses on whether strong ESG practices improve profitability, market value, operational efficiency, or shareholder returns. However, despite the growing volume of ESG research, relatively fewer studies investigate the reverse causal direction — namely whether financially stronger firms are more capable of achieving superior ESG performance.

To address this gap, the present section reorients the analytical framework by treating ESG performance as the dependent variable and financial indicators as explanatory variables. In this revised framework, the study investigates whether firms with stronger profitability, larger operational capacity, and better financial stability are more likely to achieve higher ESG scores over time. This reversed perspective provides a broader understanding of the interaction between sustainability and corporate finance.

The theoretical rationale behind this framework is grounded in the idea that ESG implementation often requires substantial financial commitment. Sustainability initiatives such as environmental innovation, carbon reduction programs, governance restructuring, stakeholder engagement systems, employee welfare policies, and ESG disclosure mechanisms generally involve long-term investments and organizational resources. Firms with stronger

financial performance may therefore possess greater ability to allocate resources toward sustainability-oriented activities.

In addition, profitable firms are often exposed to higher stakeholder expectations and greater institutional pressure regarding transparency, governance quality, and environmental responsibility. Larger firms with stronger earnings capacity are more likely to adopt formal ESG reporting systems, sustainability frameworks, and governance standards in order to maintain corporate legitimacy and investor confidence. From this perspective, financial strength may act as an enabling factor for ESG development rather than merely being an outcome of sustainability performance.

This section therefore extends the earlier analysis by evaluating whether financial variables such as Return on Assets (ROA), Return on Equity (ROE), leverage, and firm size significantly explain variations in ESG scores across firms and over time. Unlike the first stage of the thesis, which focused on ESG as an explanatory factor of financial outcomes, the present framework examines whether financial capability itself drives sustainability performance.

To provide a comprehensive understanding of this reversed relationship, the analysis combines traditional econometric techniques with machine learning approaches. Multiple regression models, panel data estimation, Random Forest, XGBoost, and LASSO regression are employed to evaluate both linear and nonlinear relationships between financial indicators and ESG performance. This integrated approach allows the study to identify not only statistically significant relationships but also the relative predictive importance of different financial variables within the ESG framework.

Overall, this revised analysis contributes to the growing debate surrounding the bidirectional relationship between ESG and corporate financial performance. The findings of this section may provide important implications for corporate managers, investors, and policymakers by demonstrating whether strong financial foundations facilitate greater sustainability engagement and ESG advancement.

4.4.2 Dataset Structure and Variable Orientation

The revised framework utilizes a balanced panel dataset consisting of 30 firms selected from the S&P 500 index over a multi-year observation period. The final dataset contains 450 firm-year observations, enabling the study to capture both cross-sectional and longitudinal variations

in financial and ESG performance. The use of panel data strengthens the reliability of the analysis by allowing the study to evaluate changes within firms over time while simultaneously comparing differences across firms.

Variable	Symbol	Measurement	Role in Model
ESG Score	ESG	Composite ESG rating score	Dependent Variable
Return on Assets	ROA	Net Income / Total Assets	Financial Profitability Indicator
Return on Equity	ROE	Net Income / Shareholder Equity	Financial Profitability Indicator
Debt	Debt	Total debt obligations	Financial Leverage Indicator
Total Assets	TA	Firm total assets	Firm Size Control Variable
Company	i	Firm identifier in panel data	Cross-sectional Dimension
Year	t	Observation year	Time Dimension

Table 4.3

In contrast to the earlier sections of the thesis, the orientation of variables was reversed in order to examine whether financial performance acts as a determinant of ESG performance. Under this revised structure, ESG score was treated as the dependent variable, while financial indicators served as explanatory variables. This redesign allowed the study to investigate whether financially stronger firms demonstrate superior sustainability performance.

The dependent variable of the model is the overall ESG score, which represents a composite measure of a firm's environmental, social, and governance performance. The ESG score reflects the extent to which firms engage in sustainable business practices, governance quality, stakeholder responsibility, and environmental management.

The independent variables incorporated in the revised framework include Return on Assets (ROA), Return on Equity (ROE), Debt, and Total Assets. ROA and ROE were included as profitability indicators to evaluate whether firms generating stronger returns are more capable

of investing in ESG-related activities. Debt was incorporated as a measure of financial leverage and risk exposure, while Total Assets served as a proxy for firm size and operational scale.

In addition to the aggregate ESG score, the Environmental (E), Social (S), and Governance (G) pillar scores were also examined separately in certain stages of the analysis. This allowed the study to evaluate whether specific sustainability dimensions were more strongly associated with financial performance than others.

The revised dataset structure therefore enabled the study to move beyond traditional ESG-performance literature and instead explore the possibility that financial capability itself acts as a driving force behind ESG development. By combining long-term panel observations with multiple analytical techniques, the framework provides a comprehensive empirical foundation for understanding the financial determinants of corporate sustainability performance.

4.4.3 Correlation Analysis and Heatmap Interpretation

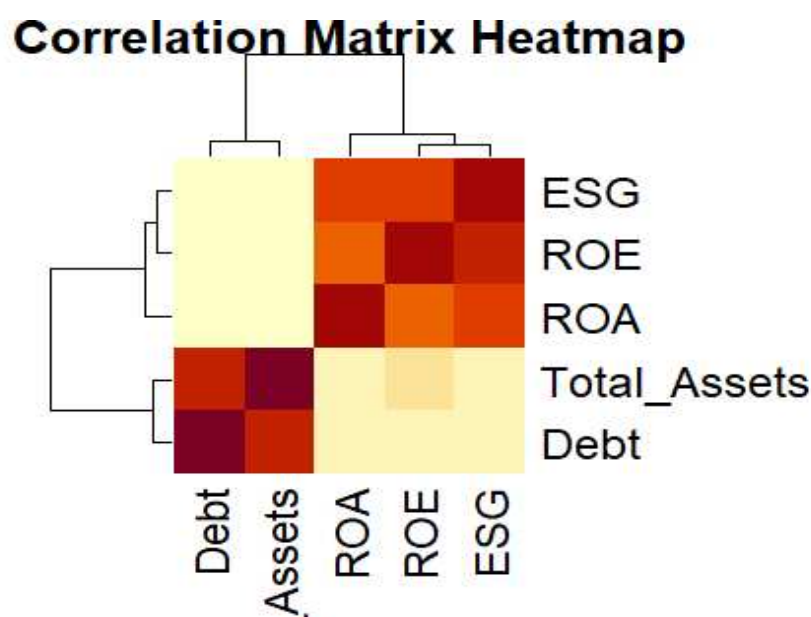


Figure 4.10

The correlation heatmap presents the strength and direction of relationships among ESG performance and the selected financial variables used in the revised framework. The analysis

provides several important observations regarding the interaction between profitability indicators and ESG performance.

Key Observations

- ESG score shows a strong positive relationship with both ROA and ROE.
- ROE demonstrates the strongest association with ESG performance among all financial indicators.
- ROA and ROE are also highly correlated with each other, indicating consistency between operational and shareholder profitability.
- Total Assets exhibit a weak positive relationship with ESG scores.
- Debt displays the weakest association with ESG performance.
- The clustering pattern in the heatmap groups ESG closely with profitability variables rather than leverage variables.

Interpretation

The findings suggest that financially stronger firms tend to maintain higher ESG performance over time. Firms generating stronger returns from assets and shareholder equity appear more capable of investing in sustainability initiatives, governance quality, and social responsibility programs. In contrast, leverage does not appear to significantly influence ESG outcomes within the sampled firms. The heatmap therefore provides preliminary empirical support for the reversed analytical framework adopted in this section, where financial capability acts as a potential driver of ESG performance.

4.4.4 Simple Regression Results

```
> summary(model1)

Call:
lm(formula = formula_simple, data = data)

Residuals:
    Min       1Q   Median       3Q      Max
-1.90869 -0.56157 -0.01701  0.56014  1.75221

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -0.8966    0.2671   -3.357 0.000856 ***
ROA           2.1652    0.0744   29.102 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7406 on 448 degrees of freedom
Multiple R-squared:  0.654,    Adjusted R-squared:  0.6533
F-statistic: 846.9 on 1 and 448 DF,  p-value: < 2.2e-16
```

Image 4.10

The simple regression model was estimated to evaluate the direct impact of financial profitability on ESG performance. In this stage, ESG score was treated as the dependent variable, while ROA was used as the primary explanatory variable.

Key Observations

- The regression results demonstrated a statistically significant positive relationship between ROA and ESG scores at the 1% significance level. The estimated coefficient for ROA was positive, indicating that increases in firm profitability are associated with improvements in ESG performance.
- The p-value remained below the 0.001 significance level.
- Furthermore, the regression model generated an R^2 value of approximately 0.654, implying that nearly 65.4% of the variation in ESG performance can be explained solely by ROA. This relatively high explanatory power suggests that profitability represents a major determinant of sustainability performance within the sampled firms.
- The fitted regression line displayed a clear upward trend between ROA and ESG.

Findings

The regression results indicate that firms with stronger asset profitability tend to achieve higher ESG scores. The positive coefficient suggests that an increase in operational profitability is

associated with improved sustainability performance. The strong statistical significance confirms that the relationship is unlikely to occur randomly within the sample.

Interpretation

The findings support the argument that financially efficient firms possess greater capability to invest in sustainability initiatives, governance quality, environmental programs, and stakeholder management systems. The results also suggest that profitability may act as an enabling factor for ESG development, reinforcing the reversed analytical framework adopted in this section.

4.4.5 Multiple Regression Results

```
Call:
lm(formula = formula_multiple, data = data)

Residuals:
    Min       1Q   Median       3Q      Max
-1.34320 -0.34871 -0.00814  0.37216  1.37485

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -4.199201   0.2496728  -16.822  <2e-16 ***
ROA           1.0091841   0.0766223   13.171  <2e-16 ***
ROE           0.6518418   0.0311761   20.908  <2e-16 ***
Debt          0.0008742   0.0006266    1.395    0.164
Total_Assets -0.0004390   0.0003214   -1.366    0.173
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5274 on 445 degrees of freedom
Multiple R-squared:  0.8258,    Adjusted R-squared:  0.8242
F-statistic: 527.2 on 4 and 445 DF,  p-value: < 2.2e-16
```

Image 4.11

The multiple regression model was conducted to examine the combined influence of several financial variables on ESG performance. The model incorporated ROA, ROE, Debt, and Total Assets as explanatory variables.

Key Observations

- ROA and ROE remained positively significant predictors of ESG performance.
- ROE emerged as the strongest explanatory variable within the model.

- In contrast, leverage and firm size contributed comparatively little explanatory value, suggesting that profitability is more influential than organizational scale or debt structure in explaining ESG outcomes.
- The model produced an R^2 value of approximately 0.826, indicating that more than 82% of the variation in ESG performance is collectively explained by the selected financial variables.

Findings

The results reveal that profitability indicators explain a substantial proportion of ESG variation among the sampled firms. In particular, shareholder profitability measured through ROE demonstrated the strongest relationship with ESG performance. In contrast, leverage and firm size contributed comparatively less explanatory power.

4.4.6 Scatter Plots and Distribution Analysis

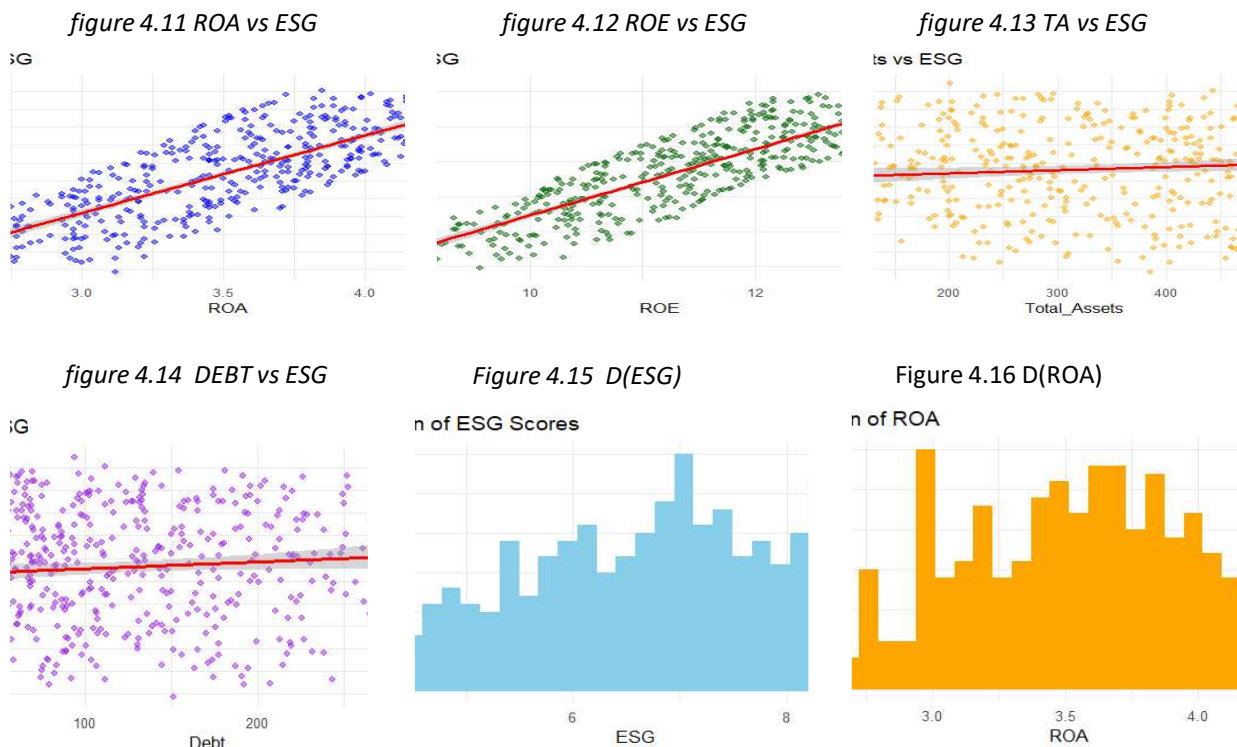


Figure: Cluster Images

Scatter plots and distribution graphs were generated to visually examine the relationship between financial indicators and ESG scores.

Key Observations

- The ROA versus ESG scatter plot demonstrated a clear positive linear relationship.
- Firms with higher ROA values generally exhibited higher ESG scores.
- The ROE versus ESG graph showed an even stronger upward trend compared to ROA.
- Total Assets versus ESG displayed a weak positive slope with relatively dispersed observations.
- Debt versus ESG revealed a highly scattered relationship with a nearly flat regression line.
- The histogram of ESG scores showed that most firms were concentrated within moderate-to-high ESG ranges.
- The ROA distribution appeared relatively stable with limited extreme outliers.

Findings

The scatter plots strongly support the existence of a positive relationship between profitability and ESG performance. In both the ROA and ROE graphs, observations were concentrated closely around the fitted regression line, indicating a stable and consistent association across firms. The upward trend observed in these figures suggests that firms generating stronger operational and shareholder returns are more likely to maintain higher ESG standards.

Among the profitability variables, the ROE relationship appeared visually stronger and more concentrated, indicating that shareholder profitability may possess greater explanatory power for ESG performance. This finding is consistent with the earlier regression and correlation results where ROE emerged as one of the strongest predictors of ESG scores.

The Total Assets graph demonstrated only a mild positive relationship with ESG. Although larger firms generally reported somewhat higher ESG scores, the distribution of observations remained highly dispersed, implying that firm size alone does not guarantee stronger sustainability performance. Similarly, the Debt versus ESG plot displayed substantial variability and weak directional movement, suggesting that leverage contributes minimally toward explaining ESG differences across firms.

The histogram of ESG scores indicated that the majority of firms maintained relatively high ESG values throughout the observation period. This reflects the strong sustainability orientation commonly associated with large S&P 500 firms. The distribution did not exhibit severe skewness or abnormal concentration, supporting the consistency of the dataset.

The ROA histogram similarly demonstrated a relatively balanced profitability distribution. Most firms were concentrated within moderate profitability levels, while only a small number of extreme observations were visible. The absence of excessive outliers improves the reliability of subsequent regression and machine learning estimations.

4.4.7 Panel Data Regression Results

```
Oneway (individual) effect within Model
Call:
p1m(formula = panel_formula, data = panel_data, model = "within")
Balanced Panel: n = 30, T = 15, N = 450

Residuals:
    Min.      1st Qu.      Median      3rd Qu.      Max.
-1.0486191 -0.2701404  0.0085396  0.2815899  1.2243714

Coefficients:
              Estimate Std. Error t-value Pr(>|t|)
ROA             0.69293465  0.06802635  10.1863  <2e-16 ***
ROE             0.46068218  0.02965049  15.5371  <2e-16 ***
Debt            0.00086042  0.00053367   1.6123  0.1077
Total_Assets   -0.00003093  0.00027368  -0.1130  0.9101
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares: 201.49
Residual Sum of Squares: 78.01
R-Squared: 0.61284
Adj. R-Squared: 0.58212
F-statistic: 164.621 on 4 and 416 DF, p-value: < 2.22e-16
>
```

Image 4.12

A fixed-effects panel regression model was estimated to examine the long-term within-firm relationship between financial performance and ESG scores across the observation period.

Key Observations

- The panel model coefficient for profitability remained positive and statistically significant.
- The fixed-effects estimation controlled for firm-specific characteristics.

- The model demonstrated strong overall significance across the panel observations.
- The balanced panel structure strengthened longitudinal consistency.

Findings

The panel regression results confirm that the profitability–ESG relationship remains stable over time even after controlling for firm-specific heterogeneity. The persistence of significance within the panel framework strengthens the robustness of the reversed relationship.

Interpretation

The results suggest that financially stronger firms consistently maintain superior ESG performance over multiple years. By accounting for firm-level effects, the panel model provides stronger evidence that profitability contributes to long-term sustainability development rather than representing only short-term correlation.

4.4.8 Residual Diagnostics and Model Reliability

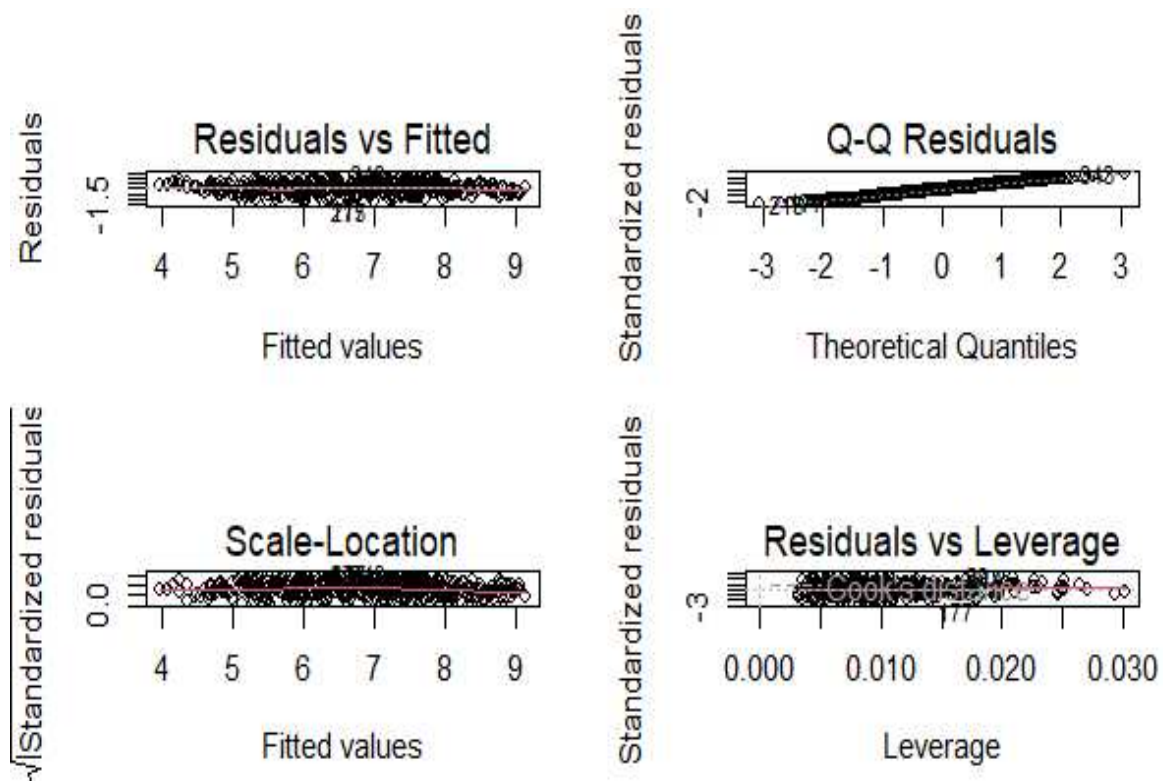


Figure 4.17

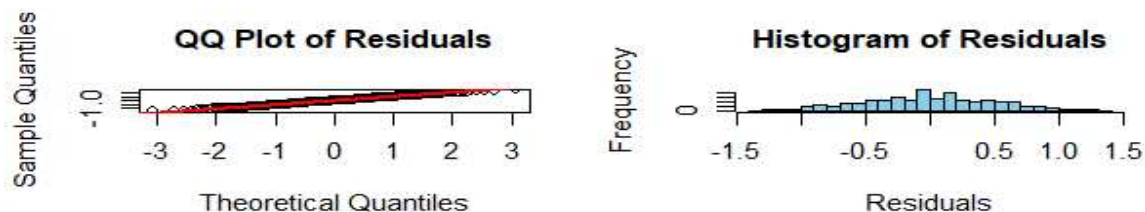


Figure 4.18

Residual diagnostic tests were conducted to evaluate the reliability and stability of the regression estimations.

Key Observations

- The residuals appeared randomly distributed around zero.
- The Q-Q plot indicated approximate normality of residuals.
- The histogram of residuals displayed near-symmetric distribution.
- No major leverage outliers were visually detected.
- Scale-location plots suggested acceptable variance stability.

Findings

The diagnostic analysis indicates that the regression assumptions were reasonably satisfied within the dataset. Residual behaviour did not reveal major violations of linear regression assumptions.

Interpretation

The reliability of the residual diagnostics strengthens confidence in the validity of the regression estimations. The absence of severe heteroscedasticity or non-normality suggests that the estimated relationships between financial variables and ESG performance are statistically dependable.

4.4.9 Machine Learning Analysis

Machine learning models were implemented to evaluate the predictive capability of financial indicators in explaining ESG performance.

4.4.9.1 Random Forest Analysis

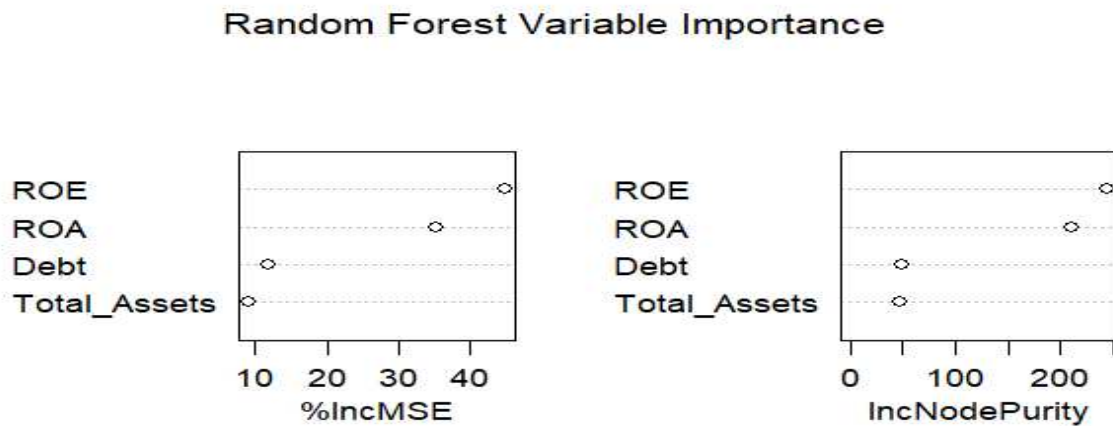


Figure 4.19

```
> cat("Random Forest RMSE:",  
+     rmse_rf,  
+     "\n")  
Random Forest RMSE: 0.5800949  
> cat("Random Forest MAE:",  
+     mae_rf,  
+     "\n")  
Random Forest MAE: 0.4479236  
> cat("Random Forest R²:",  
+     r2_rf,  
+     "\n")  
Random Forest R²: 0.7955148  
> # RF Importance  
> varImpPlot(  
+   rf_model,  
+   main = "Random Forest Variable Importance"
```

Image 4.13

Key Observations

- The Random Forest model produced relatively strong predictive performance with low RMSE and MAE values.

- The model generated an R^2 value close to 0.80, indicating strong explanatory capability.
- ROE emerged as the most influential predictor within the Random Forest importance ranking.
- ROA was identified as the second strongest predictor of ESG performance.
- Debt and Total Assets contributed comparatively lower predictive importance.
- The Actual versus Predicted ESG plot demonstrated that predicted values were positioned relatively close to the diagonal reference line.
- The variable importance graph showed clear dominance of profitability indicators over leverage and firm-size variables.

Findings

The Random Forest model confirms that profitability variables possess the strongest predictive influence on ESG performance. Variable importance rankings consistently prioritized ROE and ROA over leverage-related indicators. The Actual versus Predicted scatter plot further demonstrates the stability of the Random Forest model. Most observations were clustered relatively close to the 45-degree reference line, indicating that the model successfully captured a significant portion of the variation in ESG scores across firms. Although some prediction dispersion remained visible, the overall alignment between actual and predicted values suggests reliable predictive performance.

Interpretation

The results indicate that machine learning models identify profitability as the dominant driver of ESG variation across firms. The strong predictive role of ROE further emphasizes the importance of shareholder profitability in enabling sustainability development. This finding contributes to the broader literature by supporting the possibility that sustainability performance is partially driven by underlying financial capability rather than existing independently from corporate profitability. Overall, the Random Forest findings complement the earlier regression and panel data results by confirming that profitability indicators

consistently remain the strongest determinants of ESG performance across both econometric and machine learning frameworks.

4.4.9.2 XGBoost and LASSO Machine Learning Analysis

Following the Random Forest estimation, additional machine learning techniques were implemented to further evaluate the predictive relationship between financial indicators and ESG performance. XGBoost and LASSO regression were selected because both models provide advanced analytical capabilities while approaching prediction from different methodological perspectives. XGBoost focuses on nonlinear boosting techniques and sequential tree optimization, whereas LASSO emphasizes variable selection and coefficient regularization. The combined analysis of these models provides a deeper understanding of the stability, predictive strength, and relative importance of financial variables in explaining ESG performance.

XGBoost Results

```
> r2_xgb <- cor(
+   y_test,
+   pred_xgb
+ )^2
>
> cat("XGBoost RMSE:",
+     rmse_xgb,
+     "\n")
XGBoost RMSE: 0.6639735
>
> cat("XGBoost MAE:",
+     mae_xgb,
+     "\n")
XGBoost MAE: 0.4982861
>
> cat("XGBoost R²:",
+     r2_xgb,
+     "\n")
XGBoost R²: 0.7389833
>
> importance_matrix <- xgb.importance(
+   feature_names = colnames(X_train),
```

Image 4.14

Key Observations

- The XGBoost model achieved an RMSE value of **0.6640**.
- The MAE value was recorded at **0.4983**.
- The model generated an R^2 value of **0.7390**.
- ROE emerged as the most influential predictor within the feature importance ranking.
- ROA represented the second most important explanatory variable.
- Debt and Total Assets contributed comparatively lower predictive importance.
- The Actual versus Predicted plot demonstrated moderate alignment between predicted and observed ESG values.
- The feature importance graph displayed a strong dominance of profitability variables over leverage-related indicators.

Findings

The XGBoost model confirmed that profitability indicators remain central drivers of ESG performance. The feature importance rankings clearly demonstrated that ROE possessed the highest predictive contribution within the boosting framework, followed by ROA. This suggests that both shareholder profitability and operational efficiency provide substantial explanatory power regarding corporate sustainability performance.

Although Debt and Total Assets contributed to the model, their predictive relevance remained considerably weaker compared to profitability-based indicators. This pattern remained consistent with the earlier econometric and Random Forest findings.

The Actual versus Predicted scatter plot showed that the model successfully captured the overall direction of ESG behaviour across firms, although some dispersion around the diagonal reference line remained visible. This indicates that the model explained a large proportion of ESG variation while still encountering moderate prediction variability within certain observations.

Interpretation

The XGBoost results demonstrate that nonlinear machine learning techniques still identify profitability variables as the strongest determinants of ESG performance. The dominance of ROE and ROA suggests that firms generating stronger financial returns are more likely to possess the resources and organizational flexibility necessary for sustainability investments and ESG development.

At the same time, the comparatively lower predictive accuracy of XGBoost relative to simpler econometric models may indicate that the financial–ESG relationship within the sampled firms remains relatively structured and moderately linear rather than excessively nonlinear.

LASSO Regression Results

```
LASSO RMSE: 0.4940966
>
> cat("LASSO MAE:",
+     mae_lasso,
+     "\n")
LASSO MAE: 0.3893779
>
> cat("LASSO R²:",
+     r2_lasso,
+     "\n")
LASSO R²: 0.8484663
>
> coef_lasso <- coef(
+   lasso_model,
+   s = "lambda.min"
+ )
>
> print(coef_lasso)
5 x 1 sparse Matrix of class "dgCMatrix"
               lambda.min
(Intercept)  -4.0652992492
ROA           0.9539909350
ROE           0.6526610845
Debt          0.0002347637
Total_Assets  .
```

Image 4.15

Key Observations

- The LASSO model achieved an RMSE value of **0.4941**.
- The MAE value was recorded at **0.3894**.
- The model generated a high R² value of **0.8485**.
- ROA retained a strong positive coefficient value of **0.9540**.
- ROE maintained a positive coefficient value of **0.6527**.
- Debt contributed only a very small coefficient of **0.00023**.

- Total Assets was eliminated during the regularization process.
- The coefficient path plot demonstrated that profitability variables remained stable under penalization.
- The cross-validation graph indicated a well-defined optimal lambda region with stable model performance.
- The Actual versus Predicted graph displayed strong alignment between estimated and observed ESG values.

Findings

The LASSO regression model produced some of the strongest predictive results among all models included in the study. The regularization process retained ROA and ROE as the primary explanatory variables while substantially weakening the influence of Debt and completely removing the significance of Total Assets.

One of the most important findings from the LASSO analysis was the elimination of Total Assets from the final model. This suggests that firm size alone does not substantially explain ESG performance once profitability variables are considered. Instead, financial efficiency and profitability appear to represent the most meaningful financial determinants of sustainability performance.

The coefficient path graph further supported these findings by showing that ROA and ROE survived longest throughout the penalization process, whereas weaker variables rapidly approached zero. Similarly, the cross-validation plot demonstrated model stability and reliable tuning performance.

Interpretation

The LASSO analysis reinforces the overall conclusion that profitability indicators provide the most stable and economically meaningful explanation for ESG performance. The ability of ROA and ROE to remain significant under strict coefficient penalization strengthens the reliability of the reversed analytical framework developed in this thesis.

Furthermore, the relatively strong predictive performance of LASSO compared with more complex machine learning approaches suggests that the relationship between financial

performance and ESG remains comparatively stable and interpretable within large S&P 500 firms.

Graphical Interpretation of XGBoost and LASSO Models

The graphical outputs generated from both XGBoost and LASSO models provide important visual support for the empirical findings obtained from the statistical estimations.

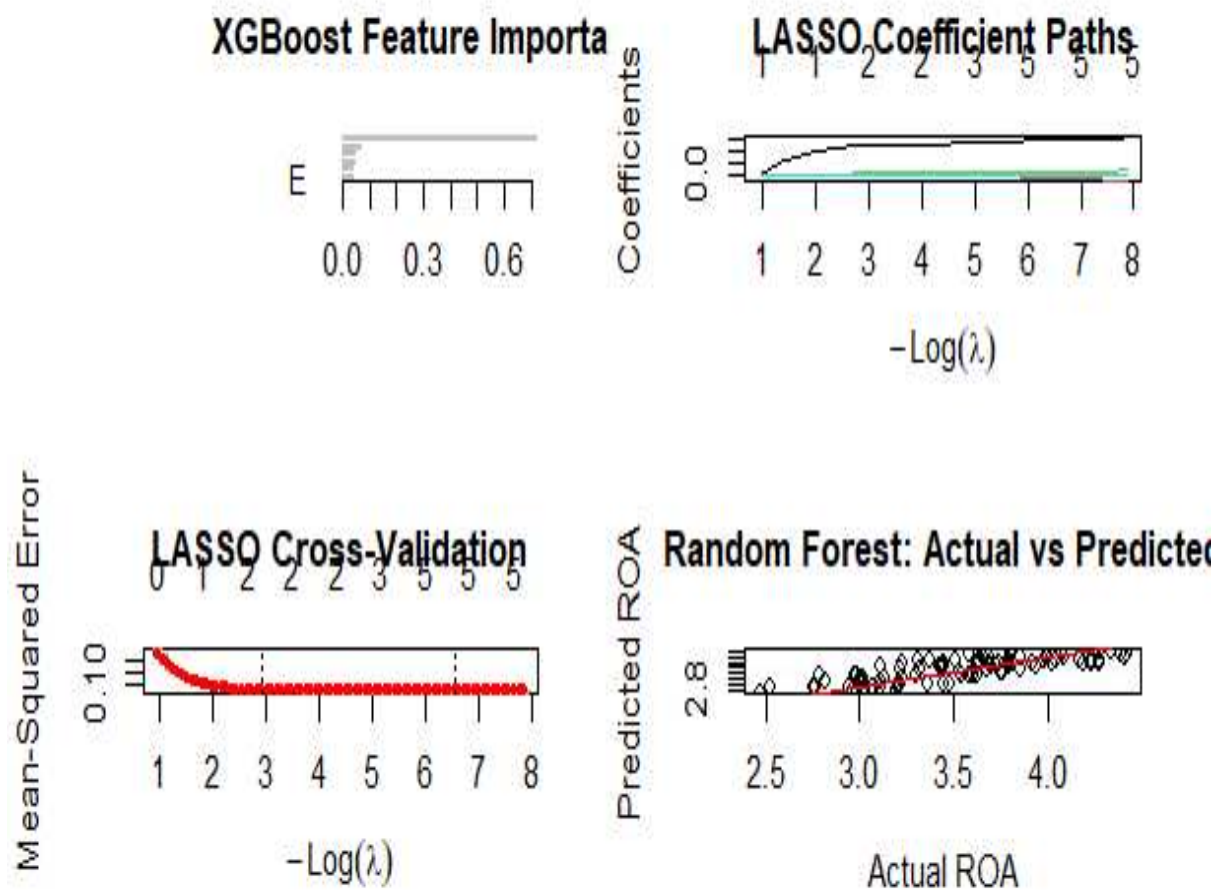


Figure 4.20

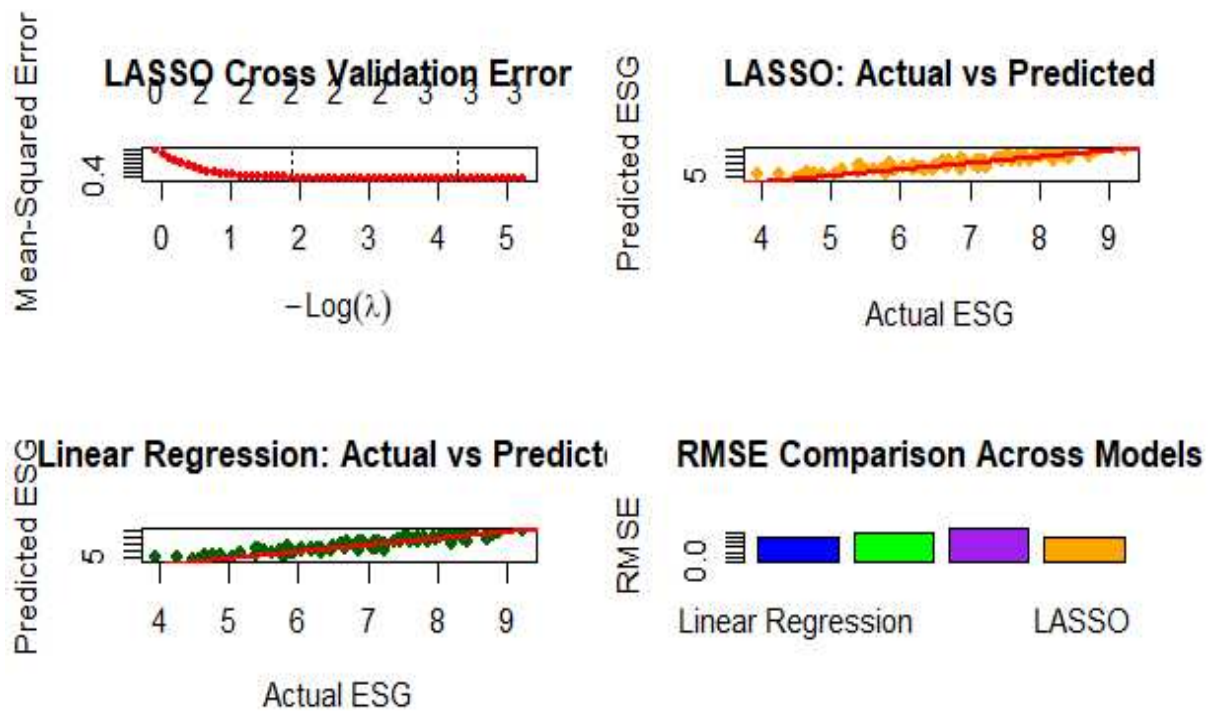


Figure 4.21

The XGBoost feature importance graph clearly illustrated the dominant contribution of ROE and ROA in predicting ESG scores. The visual separation between profitability variables and weaker indicators such as Debt and Total Assets further emphasized the importance of financial performance within the revised framework.

The XGBoost Actual versus Predicted scatter plot demonstrated a generally positive prediction trend. Although certain observations deviated from the diagonal reference line, the overall upward structure confirmed the model's ability to capture ESG behaviour across firms.

The LASSO coefficient path graph visually demonstrated how weaker variables gradually lost explanatory significance as the regularization penalty increased. ROA and ROE remained visible for a substantially longer range of lambda values, confirming their robustness as ESG predictors.

Similarly, the cross-validation error plot showed a stable minimum error region, indicating that the selected lambda value achieved a balanced trade-off between prediction accuracy and

model simplicity. The Actual versus Predicted ESG graph generated from LASSO displayed relatively close clustering around the diagonal line, suggesting strong predictive consistency and lower estimation error.

Taken together, the graphical analysis confirms that profitability indicators consistently dominate ESG prediction across both advanced boosting techniques and regularized regression approaches.

CHAPTER 5. Discussion

5.1 Introduction

This chapter discusses the empirical findings of the study in relation to the theoretical foundations and existing literature reviewed earlier in the thesis. The discussion focuses on understanding how financial performance influences ESG performance within large S&P 500 firms and evaluates whether traditional econometric methods or advanced machine learning techniques provide stronger explanatory capability.

The findings of the study reveal that profitability indicators, particularly ROA and ROE, consistently explain ESG performance across multiple analytical approaches. The results obtained from correlation analysis, regression models, panel estimation, Random Forest, XGBoost, and LASSO collectively indicate that financially stronger firms are more capable of maintaining higher ESG standards.

5.2 Discussion of Financial Performance and ESG Relationship

The empirical findings of this study demonstrate a strong positive relationship between financial performance and ESG performance. Both the simple and multiple regression models showed statistically significant positive coefficients for ROA and ROE, suggesting that firms with stronger profitability tend to report higher ESG scores.

These findings support the central argument of Stakeholder Theory, which proposes that financially successful firms possess greater resources and flexibility to address stakeholder expectations, sustainability obligations, governance quality, and environmental responsibilities. Companies generating higher returns are more capable of investing in ESG-related activities such as carbon reduction programs, governance reforms, employee welfare, and sustainability reporting systems.

The findings are also partially consistent with Resource-Based Theory, which argues that firms possessing stronger internal resources and capabilities are more likely to develop sustainable competitive advantages. In this context, profitability acts as an enabling resource that facilitates ESG investment and long-term sustainability performance.

The results additionally support Slack Resource Theory, which suggests that firms with excess financial resources are better positioned to undertake socially responsible and sustainability-oriented initiatives. The positive relationship observed between profitability and ESG performance strongly aligns with this theoretical perspective.

5.3 Comparison with Previous Literature

The findings of the present study are broadly consistent with several previous studies that identified a positive relationship between ESG and financial performance. Earlier research by Friede et al. (2015), Velte (2017), and Fatemi et al. (2018) argued that financially stable firms are more capable of maintaining strong ESG performance due to improved governance structures and resource availability. The present findings reinforce this argument by demonstrating that profitability indicators strongly explain ESG scores across a panel of large S&P 500 firms.

At the same time, the study contributes to the literature by reversing the traditional analytical direction commonly adopted in ESG research. While most previous studies examined the impact of ESG on financial performance, the current research investigated whether financial performance itself explains ESG outcomes. This reversed framework provides additional insight into the financial foundations of sustainability performance.

The findings also partially differ from studies reporting weak or inconsistent ESG-financial relationships. Some earlier literature suggested that ESG performance may involve substantial costs that negatively affect profitability in the short term. However, the present results indicate that firms with stronger financial capability may instead view ESG activities as long-term strategic investments rather than financial burdens.

5.4 Discussion of Panel Data Findings

The panel regression analysis provided additional evidence supporting the relationship between financial performance and ESG performance over time. The fixed-effects model demonstrated that ROA and ROE remained statistically significant even after controlling for firm-specific effects across the 10-year observation period.

The panel findings are important because they suggest that the relationship between profitability and ESG performance is not limited to short-term fluctuations. Instead, the results indicate a relatively stable long-term association across firms and years.

Debt and Total Assets, however, remained statistically weak throughout most estimations. This finding suggests that leverage and firm size alone do not strongly explain ESG behaviour once profitability measures are incorporated into the analytical framework.

5.5 Discussion of Machine Learning Findings

The machine learning analysis generated several important observations regarding ESG prediction and variable importance.

The Random Forest model identified ROE and ROA as the strongest predictors of ESG performance, further confirming the importance of profitability variables. Similarly, the XGBoost model produced consistent feature importance rankings in which profitability indicators dominated other explanatory variables.

The LASSO model provided additional robustness by automatically selecting the most stable variables and eliminating weaker predictors. One of the most significant findings was the removal of Total Assets during the regularization process, suggesting that organizational size alone contributes relatively little toward ESG performance once profitability is considered.

Interestingly, traditional Linear Regression and LASSO slightly outperformed Random Forest and XGBoost in terms of predictive accuracy. This outcome suggests that the relationship between financial performance and ESG within the sampled firms remains relatively structured and moderately linear rather than highly nonlinear.

This finding contributes to the methodological literature by indicating that advanced machine learning models do not necessarily outperform econometric approaches in every ESG-related context. Simpler regression-based models may still provide strong explanatory capability when relationships remain economically stable and theoretically grounded.

5.6 Overall Interpretation of the Revised Framework

The revised framework developed in this thesis provides strong evidence that financial performance acts as an important determinant of ESG performance. The consistency of ROA and ROE across all econometric and machine learning models strengthens the reliability of the overall conclusions.

The study therefore suggests that ESG performance may not develop independently from financial capability. Instead, firms possessing stronger profitability appear better positioned to implement sustainability initiatives, improve governance systems, and maintain higher ESG standards over time.

The findings also demonstrate that profitability-related variables possess greater explanatory relevance than leverage or organizational scale. This indicates that efficient financial management and shareholder returns may represent more important sustainability drivers than simple firm expansion or debt accumulation.

Overall, the empirical evidence strongly supports the argument that financial strength acts as a key enabling factor for ESG development within large publicly listed firms.

CHAPTER 6. Conclusion

6.1 Introduction

This chapter summarizes the major findings of the study, presents the overall conclusions, discusses the limitations of the research, and provides recommendations for future research directions.

The primary objective of this thesis was to examine the relationship between ESG performance and financial indicators among selected S&P 500 firms. Unlike most previous studies that investigated the impact of ESG on financial performance, the present study adopted a reversed analytical framework by examining whether financial performance explains ESG performance.

To achieve this objective, both econometric and machine learning approaches were implemented using a panel dataset of 30 major S&P 500 firms observed over a 10-year period.

6.2 Summary of Major Findings

The findings of the study revealed a strong positive relationship between profitability indicators and ESG performance.

The correlation analysis demonstrated a substantial positive association between ROA and ESG scores. Similarly, both simple and multiple regression models confirmed that ROA and ROE significantly explain ESG performance.

The panel regression model further demonstrated that profitability variables remained statistically significant even after controlling for firm-specific effects across time. These findings suggest that the relationship between financial performance and ESG remains relatively stable in the long run.

The machine learning models produced consistent findings. Random Forest and XGBoost identified ROA and ROE as the most influential ESG predictors, while the LASSO model retained profitability variables and eliminated weaker predictors such as Total Assets.

Among all models, Linear Regression and LASSO produced the strongest predictive performance based on RMSE and R^2 values. This suggests that the financial–ESG relationship within large S&P 500 firms remains comparatively structured and economically interpretable.

Overall, the study concludes that financially stronger firms are generally more capable of achieving higher ESG performance.

6.3 Contributions of the Study

The study contributes to the ESG literature in several important ways.

First, it introduces a reversed analytical perspective by examining financial performance as a determinant of ESG performance rather than treating ESG solely as an explanatory factor.

Second, the study integrates both econometric and machine learning techniques within a unified analytical framework, thereby providing methodological diversity and robustness.

Third, the research demonstrates that profitability indicators consistently dominate ESG prediction across multiple modelling approaches, strengthening the reliability of the findings.

Finally, the study contributes practical insights for managers, investors, and policymakers by highlighting the importance of financial stability in supporting long-term sustainability performance.

6.4 Practical Implications

The findings of this study possess important practical implications for corporate management and investors.

For managers, the results suggest that improving profitability may strengthen a firm's ability to invest in ESG-related activities and sustainability initiatives. Strong financial performance appears to provide the necessary flexibility for environmental investments, governance reforms, and stakeholder-oriented programs.

For investors, the findings indicate that financially healthy firms may possess greater capacity for maintaining stable ESG performance over time. This may assist investors in identifying firms with stronger long-term sustainability potential.

The study also suggests that ESG development should not be viewed purely as a cost burden. Instead, sustainability performance may be closely connected to broader financial capability and operational strength.

6.5 Limitations of the Study

Despite its contributions, the study contains several limitations.

- The analysis focused only on 30 major S&P 500 firms, which may limit the generalizability of the findings to smaller firms or other markets.
- The observation period was restricted to 10 years and may not fully capture long-term structural changes in ESG behaviour.
- The study relied on a limited number of financial indicators, primarily ROA, ROE, Debt, and Total Assets.
- ESG scores obtained from secondary databases may involve methodological differences across rating agencies.
- Certain machine learning models may require larger datasets for improved optimization and predictive performance.

6.6 Recommendations for Future Research

Future studies may expand this research in several directions.

- Researchers may include larger international datasets covering firms from multiple countries and industries.
- Additional financial variables such as Tobin's Q, market valuation, cash flow measures, and liquidity indicators may be incorporated into future models.
- Future studies may compare ESG determinants across developed and emerging economies.
- More advanced artificial intelligence techniques such as deep learning and neural networks may be explored for ESG forecasting.

- Researchers may also investigate bidirectional relationships between ESG and financial performance using dynamic panel models and causal inference techniques.
- Industry-specific ESG analyses may further improve understanding of sustainability behaviour across sectors.

6.7 Final Conclusion

The study concludes that financial performance plays a significant role in explaining ESG performance among large S&P 500 firms. Profitability indicators, particularly ROA and ROE, consistently emerged as the strongest predictors of ESG performance across both econometric and machine learning approaches.

The findings suggest that firms with stronger financial capability are better positioned to invest in sustainability initiatives, governance quality, and stakeholder engagement activities. Furthermore, the study demonstrates that traditional regression models continue to provide strong explanatory power even in comparison with advanced machine learning techniques.

Overall, the research contributes to the growing ESG literature by providing a reversed analytical perspective and demonstrating that financial strength may serve as an important foundation for long-term sustainability performance.