ERASMUS UNIVERSITY ROTTERDAM
ERASMUS SCHOOL OF ECONOMICS
MSc Economics & Business
Specialization Financial Economics

Exploring the Impact of ESG Scores on the Financial Performance of Firms

Fixed Effects Analysis of ESG Scores Reported for 2018-2021

Author: Martijn van Buuren
Student number: 584112
Thesis supervisor: Soebhag, A.
Second reader: Van Vliet, B. P.
Finish date: August 2024

PREFACE AND ACKNOWLEDGEMENTS

This thesis is the results of the accumulation of academic skills that I have learned in the 4 years of studying at the Erasmus University Rotterdam. In this thesis I delved deeper into understanding the Environmental, Social, and Governance (ESG) scores and its impact on financial performance of firms. The ESG scores have been a growing interest of mine since I learned about them during one of my earlier group projects. Since then I was interested in gaining a deeper understanding of these scores by exploring the existing literature and learning how these scores are computed and what their effects are on different aspects of the firms, in particular their impact on the financial performance of the firms. This thesis was a great opportunity to explore this subject and to apply the skills I have learned and developed while studying at the Erasmus University.

I want to express my gratitude to my supervisor for the time he put into guiding me through the process of writing my thesis. His quick responses and great communication were really admirable and helpful while working on my thesis. I am also very thankful for the great and useful feedback, which helped me to spot the mistakes, which I would otherwise have missed, making this thesis even better than I could have initially imagined. Lastly I want to thank my friends and the Erasmus University for giving me the motivation, resources, and the help needed during my career at the Erasmus University to achieve all the amazing accomplishments. It has been a great honour that I had the chance to spend the last 4 years as a student at the Erasmus University and that I was able to learn from the smartest and most prestigious professors I've ever met. I wouldn't have wished it any other way.

ABSTRACT

This paper studied if higher ESG scores lead to better financial performance for firms. This relationship was studied using Ordinary Least Squares (OLS) regression models with fixed effects. The data that was used for this study was taken from two different sources. The Thomson Reuters ESG score data was provided by the Erasmus University's Data Service Centre (EDSC) and the fundamental company data was taken from the Compustat database, available at the Wharton Research Data Services (WRDS). The sample used for this study span from 2018 to 2021. The results showed no significant effect of ESG scores on the financial performance within firms when considering the full sample, however when interaction terms with the different sectors were included, the estimates showed significant coefficient estimates. These results suggest that there are differences in the impact of higher ESG scores on a firm's financial performance between sectors.

Keywords: ESG Scores, Firm Performance, Corporate Governance, Sustainability, Corporate Social Responsibility (CSR)

JEL Classification: G30, G32, L25, Q56

TABLE OF CONTENTS

CHAPTER 1 Introduction

In today's global business environment, there has been a tremendous shift and focus on global issues. This has led to a need for a framework to analyze to which extend companies are focused on solving Environmental, Social and Governance issues. This has in turn led to the so called Environmental, Social and Governance score, also known as ESG score. Especially in the last few years more companies have started to report to what extent they are focusing on and implementing ESG policy in their business operations. It is estimated by the Bloomberg Intelligence (Bloomberg, 2021) that the global ESG assets are expected to exceed $53 trillion by 2025. There could be a few reasons for the growing importance of ESG. It could be that companies want to create a positive reputation, to use as a marketing tool to attract more investments. Another reason could be that the consumers, investors, shareholders, and governments demand more transparency, accountability and sustainability. There is however still a lot of debate with regards to the effectiveness of the ESG scores on solving the global, social, environmental, and governance issues. It is also still debated whether ESG scores constitute more efficient investing and company policies. This paper delves deeper into understanding the importance of company policies with regards to ESG on the financial performance of the firm. To get a better understanding of this, it is essential to define the term ESG first.

The ESG criteria constitute a set of standards for a company's operations that socially conscious investors use to screen potential investments. Environmental criteria consider how a company performs as a steward for the planet. Social criteria examine how it manages relationships with employees, suppliers, customers. Governance deals with a company's leadership, executive pay and making sure everything's above board. The usefulness of ESG criteria comes from their ability to provide a better picture of a company's health and prospects, beyond its financial statements. Giese, et al. (2019) show in their research that changes in a company's ESG characteristics may be a useful financial indicator. A problem arises when trying to rate companies based on ESG scores according to Billio, et al. (2021). They show that there is heterogeneity in rating criteria among lead agencies evaluating companies. This leads to agencies having opposite opinions on the same evaluated companies. This is due to the highly subjective nature of the ESG criteria, which can be interpreted differently for every firm. There can be differences in the focus on certain aspects of ESG, based on for example the industry a certain firm is operating in. And there can be many other aspects which can have an impact on the evaluation of a company's ESG score, making it difficult to compare firms with one another.

Although it is difficult to assess how a firm's ESG score is precisely determined, nonetheless more firms have started to report these scores. This growing importance of the Environmental, Social, and Governance concepts as additional criteria to evaluate companies, raises the following research question:

do higher ESG scores constitute to better financial performance for firms? The relationship between ESG scores and firm performance has been extensively studied, with research findings pointing in various directions. A number of studies highlight a positive correlation, suggesting that better ESG scores are associated with better financial performance, and better stock performance. Friede et al., for example, performed a meta-analysis, which was published in 2015. Their paper was reviewed by over 2 thousand empirical studies of which a majority reported a positive relationship between ESG and financial performance, using metrics such as return on equity (ROE) and return on assets (ROA). Similarly Khan et al. (2016) showed that firms with good sustainability ratings outperformed firms with poor sustainability ratings. Amel-Zadeh & Serafeim (2018) also highlighted that ESG ratings were associated with lower cost of capital and less earnings volatility, suggesting that ESG performance might contribute to firm stability. Even though these studies show positive relationships between higher ESG scores and better firm performance, there are also studies that show an insignificant relationship or even a negative significant relationship between firm performance and higher ESG scores. For example, Barnett & Salomon (2006) have either identified no significant relationship or even a negative correlation between ESG scores and firm performance in certain contexts or industries, possibly due to the substantial costs of ESG initiatives without immediate financial returns. Some other papers support this finding and show that the relationship does not extend to market-based financial performance, as measured by Tobin's Q (Velte, 2017). Some of these studies may however already be outdated, due to the dynamic and everchanging financial environment. The diversity in findings could be attributed to differences across industries, geographical regions, the date that the study was published, and the varied methodological approaches taken by researchers, including how they measure ESG performance and financial outcomes.

Whether higher ESG scores constitute to better financial performance for firms, will be studied by first examining the effect of the overall and the combined ESG scores on the firm's financial performance. The next step will be to delve deeper into this question by differentiating between the subcategories of ESG criteria. It could be that when the overall or combined ESG scores do not show a significant relationship with the firm's financial performance, that some of the ESG measures which make up the overall and combined ESG scores do show a significant relationship. If this paper finds a positive effect of ESG score on a firm's financial performance, it could help firms understand which criteria should be focused on to enhance the firm's financial performance most efficiently. If this paper does not find a significant positive relationship between higher ESG scores and firm performance, this may indicate that the ESG scores have not shown significant results for the financial performance of the firms yet. This does not necessarily show that the ESG scores are (in)effective in exploring other measures of a firms performance, for example sustainability, or societal impact.

The methods that will be applied to answer the research question mentioned above are the following. In this paper the same independent variables that were chosen by Friede et al. (2015) will be used as proxies for firm performance, however instead of a meta-analysis, this paper will study the same relationship using an ordinary least squares regression suitable for panel data analysis. The dependent variables for the models will be return on assets and return on equity. The independent variable will be the ESG score, which will be split into a model with the overall ESG score, a model with the combined ESG score, and a model analyzing the effect of different subcategories of the ESG scores that make up the overall and combined scores. These scores can be separated into 10 different criteria categorized into the three main pillars: the environmental, social, and governance pillar. To make the models more robust in explaining the effect of ESG scores on firm performance, a number of control variables will be added to the models. Those include: country, industry, firm size, leverage ratio, etc.. These control variables can account for a number of differences in characteristics between firms. Firm size, for example, can have an effect on the ESG scores of companies, because larger firms may have more resources, for example social networks or more political influence, to be able to get higher ESG scores, compared to smaller companies. However there may also be more people paying attention to these companies, leading to more critics observing these companies, which can in turn lead to lower scores. The variable country could differentiate between geographical locations, which can account for differences in regulatory pressures between countries, cultural differences, etc.. Lastly different statistical tests will be performed to test the validity of the models and the robustness of the results.

This study requires two different kinds of data, ESG data and financial company data. The ESG scores data follows the Thomson Reuters standardized framework of reporting the ESG criteria. The scores are split into three main pillars: Environmental, Governance, and Social. These three main pillars are again divided into three to four subcategories. For the environmental score this includes: resource use, emissions, and innovation. Governance is divided into: managements, shareholders, and Corporate Social Responsibility (CSR) strategy. And lastly the social pillar is divided into: workforce, human rights, community, and product responsibility. This data is collected into two large databases by the Erasmus University's Data Service Centre. These databases contain the ESG scores for over 6,000 public companies globally for the years 2000-2023 and 2010-2022. The data for 2000-2023 has a slightly smaller universe, but contains additional static information. The 2010-2022 data contains ESG scores, but no additional static information. The second data that is needed to perform this study is financial company data that is available at the Compustat database, containing a large number of company fundamentals that publicly traded companies are required to report, for example balance sheet items. The companies in the ESG scores database are identified with an International Securities Identification Number, or ISIN code, which can be searched for in the Compustat database using the Wharton Research Data Services digital platform. The data will then be merged into one large comprehensive database linking the ESG scores and the financial data to the companies identified by their ISIN code and sorted by years. The final dataset consists of unbalanced panel data, because reporting ESG is not mandatory, and thus can be missing for some companies, and for certain years. Although it has become more important, not all companies are willing to report this, and some have only started to report their ESG performance recently. Consequently many companies will have missing data for the earlier years that are contained in the dataset. To handle the mostly missing observations for the first few years in the data, only the years 2018-2021 will be included in the sample. The rest of the paper will be structured as follows: firstly this paper will provide a literary review, secondly it will propose a methodology and the data that will be used to study the relationship, next it reports the results, and lastly it will report and discuss the findings, and conclude the paper.

CHAPTER 2 Theoretical Framework

2.1 The Relevance of ESG

As of 2018 over 6,000 global public companies have started to report their ESG scores according to a 2018 report of Thomson Reuters, using the Thomson Reuters' ESG framework. Their ESG framework constitutes a set of standards for a company's operations that socially conscious investors use to screen potential investments. It consists of three main criteria, also called pillars within the ESG framework. The first is the environmental criteria, which considers how a company performs as a steward for the planet. The social criteria examines how a firm manages relationships with employees, suppliers, customers, etc.. The criteria for governance deals with a company's leadership, executive pay and making sure everything's above board. The usefulness of these ESG criteria comes from their ability to provide a better picture of a company's health and prospects, beyond its financial statements. The three main pillars are divided into even more subcomponents that look at these criteria in more detail. These subcomponents are given a score which are then used to determine the overall ESG score for a company. The environmental pillar is subdivided into the measures: resource use, emissions, and innovation. The social pillar consists of the measures: work force, human rights, community, product, and responsibility. Lastly the governance pillar of the ESG framework consists of the measures: management, shareholders, and CSR strategy. These subcategories consist of even more sub measures. The data to determine the Thomson Reuters ESG scores up to 2018 consisted of a total of 400+ sub measures, however this study will only focus on the three main pillars and their corresponding 10 subcategories. An extensive literature and a large number of studies have investigated the effectiveness of these scores, with regards to the effects on the operational side of firms. More importantly, studies have looked at the effects of the ESG scores on the firms' corporate and financial performance. A number of these studies are reported below.

First of all, Giese, et al. (2019) show in their research that changes in a company's ESG characteristics may be a useful financial indicator. A problem arises when trying to rate companies based on ESG scores according to Billio, et al. (2021). They show that there is heterogeneity in rating criteria among lead agencies evaluating companies. This leads to agencies having opposite opinions on the same evaluated companies. This finding highlights the subjective nature of the ESG scores. Although it is difficult to assess how a firm's ESG score is precisely determined, nonetheless more firms have started to report these scores, making these scores an ever growing importance in evaluating firms. It could be taken into question if these scores are effective in increasing the financial performances of firms that are concerned with ESG. This raises the following research question: Do better ESG scores constitute to better financial performance for firms?

2.2 Findings of Previous Studies

The relationship between ESG scores and firm performance has been extensively studied, with research findings pointing in various directions. A number of studies highlight a positive correlation, suggesting that better ESG scores are associated with better financial performance, and better stock performance. Friede et al., for example, performed a meta-analysis, which was published in 2015. Their paper was reviewed by over 2 thousand empirical studies of which a majority reported a positive relationship between ESG and financial performance, using metrics such as return on equity (ROE) and return on assets (ROA). Similarly Khan et al. (2016) showed that firms with good sustainability ratings outperformed firms with poor sustainability ratings. Amel-Zadeh & Serafeim (2018) also highlighted that ESG ratings were associated with lower cost of capital and less earnings volatility, suggesting that ESG performance might contribute to firm stability. Even though these studies show positive relationships between higher ESG scores and better firm performance, there are also studies that show an insignificant relationship or even a negative significant relationship between firm performance and higher ESG scores. For example, Barnett & Salomon (2006) have either identified no significant relationship or even a negative correlation between ESG scores and firm performance in certain contexts or industries, possibly due to the substantial costs of ESG initiatives without immediate financial returns. Some other papers support this finding and show that the relationship does not extend to market-based financial performance, as measured by Tobin’s Q (Velte, 2017). Some of these studies may however already be outdated, due to the dynamic and everchanging financial environment. The diversity in findings could be attributed to differences across industries, geographical regions, the date that the study was published, and the varied methodological approaches taken by researchers, including how they measure ESG performance and financial outcomes.

More recent papers, for example Halid et al. (2023) report and review the results of 11 different studies that research the impact of ESG scores on the financial performance of listed companies. They report both positive and negative relationships between ESG scores and firm performance. The positive studies were in the majority with 7 out of the 11 studies reporting a strictly positive relationship, 2 studies reporting a strictly negative relationship, and lastly 2 studies reporting a neutral or mixed relationship between firm performance and ESG scores. The majority of these studies report a positive relationship, which gives an indication of the results that would be expected to be found when performing a similar study with more recent data. Another study done by Nguyen et al. (2022) had investigated the impact of ESG on firms’ financial performance using a sample consisting of 57 U.S. non-financial firms listed on the S&P 500. They used ROA, ROE, and Tobin’s Q as measures for firm performance. They performed a two-staged least squares regression with firm and year fixed effects. Their control variables where: leverage ratio, R&D, and firm size. They find a positive significant relationship between higher ESG scores and ROA, ROE, and Tobin’s Q. The coefficient that they estimated for Tobin’s Q is however much larger than the estimated coefficients for ROA and ROE, but all are positive and significant at the 5% confidence level.

The study concludes that having a better practice of ESG could enhance firms' financial performance as measured using the variables reported above.

Chen et al. (2023) performed study with a more elaborate dataset consisting of 3332 listed companies worldwide, over the span of 10 years (2011–2020). They used a total of 24,076 observations. They applied multiple regression and categorized regression to their sample and used only ROA as their independent variable measuring firm's financial performance. The study's findings showed that ESG performance is positively interrelated with corporate performance at the 1% confidence level. According to the findings of the study, their regression coefficient for ESG is significant at the 1% confidence level, and has a positive sign. Their results conclude that the influence of ESG rating on corporate performance is significant for large-scale companies and insignificant for small-scale companies. The influence of firm size on environmental, social and governance (ESG) ratings is examined in the study performed by Gregory (2022). The study uses a dataset consisting of 1601 international firms from 2011 to 2019. It controls for ratings agency and industrial sector. There are five ratings agencies included in the study. A uniform positive relationship is found between firm size and ESG over ratings agencies, but its strength varies by agency, calling into question the explanation of organizational legitimacy. The results show that for many combinations of ratings agency and industrial sector, there is no significant relationship between firm size and ESG rating. The results also show that the effect of size on ESG ratings is driven in part by outliers. These findings are important because ESG ratings are primarily determined by organizational legitimacy. The results indicate that when using ESG scores as regressors, it may be appropriate to control for the effects of firm size.

Priem and Gabellone (2024) looked at the relationship between the ESG score and the cost of capital of 600 companies across 17 European countries that are part of the EURO STOXX 600 index. They found that companies with higher ESG scores have a lower cost of capital. This relationship however only holds for firms located in countries with a weaker legal environment. These results indicate that the leverage ratio is an important part in understanding the impact of ESG on a firm's financial performance. Another study focused on the subcomponents of main ESG pillars (Mashayekhi et al., 2024). They looked at the importance of the ESG pillars and their subcomponents at both the firm and industry level, using the Thomson Reuters' ESG scores database. The sample they used for their study span from 2009 to 2017. A K-means cluster analysis was performed to determine the most important ESG pillars. The results showed that the social and economic pillars are the two fundamental pillars of ESG performance in all industries in general. However this can differ from industry to industry. This raises the question of whether there are differences in the effectiveness of the ESG scores on the firms' financial performance between different industries. Abhayawansa & Tyagi (2021) have noted a trend in ESG investing. It has gained momentum since the COVID-19 pandemic, with an increase in interest for ESG investing. This has in turn led to a growing demand for ESG data, ratings and rankings. This increase in demand for ESG data went together with a growing number of agencies offering this data. The problem still persists that there are differences between the ratings and rankings of these agencies, making it difficult to compare them with each other. This paper contributes to the existing literature by studying the effects of ESG scores on firms' financial performance using a large number of companies for which their ESG scores are all computed using the Thomson Reuters' ESG framework. The next section describes the data and methods used to study this relationship in more detail.

CHAPTER 3 Data

3.1 Thomson Reuters

The data that is used for this study consists of a large database of publicly traded global companies for which the ESG scores are reported and collected. The database that reports these ESG scores is provided by the Erasmus University's data service center (EDSC). The ESG scores are computed according to Thomson Reuters' strategic ESG framework (Thomson Reuters, 2018). The information needed for the calculation of the ESG scores is collected based on publicly available sources such as company websites, annual reports, and corporate social responsibility reports or contributed by firms, then audited and standardized. The data covers around 6,500 global companies. The initial data consisted of the years 2010 to 2022, however for most of the earlier years the ESG scores were missing, which is in line with the momentum in ESG data reporting that Abhayawansa & Tyagi (2021) had noticed since the COVID-19 pandemic. To deal with the missing data, the timeframe for this study is shortened to include only the years 2018 to 2021. The dataset consists of panel data, however it is highly unbalanced due to many observations missing. Some companies have started to report their ESG scores in the most recent years, and for most companies their scores for 2022 haven't been included in the data yet. In some cases the observations are not available for every consecutive year. This leaves some gaps within the data. A problem that may arise is selection bias, because it could be the case that firms that know beforehand that their scores would be bad, could opt out of having these scores. Also firms that do not report their scores for each year are excluded to prevent an unbalanced sample, which may lead to a bias that coincides with certain systematic firm characteristics for the firms that are excluded.

3.2 Compustat

The next step in gathering the necessary data is to add fundamental company data for the companies that are included in the ESG score dataset. A few examples of the fundamental company data are: balance sheet items, income statement items, and other firm characteristics. The ESG data and the fundamental data is linked based on the ISIN codes. These codes are used to identify publicly traded global companies. The fundamental company data comes from the Compustat database provided WRDS. After cleaning the data, merging the ESG score data with the Compustat data, and removing duplicates, a total of 28,853 observations remained. These observations still included missing values for the ESG scores. After dropping the observations with missing values, a total of 18,842 observations remained in the final dataset across all variables. Although not every company has reported data for each year. Only the years for which the ESG scores were reported are included in the final dataset.

3.3 ESG Score Methodology

The following methodology is used to compute the ESG scores that are used for this study. The ESG scores can be categorized into a number of different subcategories (Thomson Reuters, 2018). These subcategories are divided into three main pillars of ESG: Environmental, Social, and Governance. The scores for the subcategories are calculated based on a percentile rank scoring methodology, which is based on three factors: How many companies are worse than the current one? How many companies have the same value? And how many companies have a value at all?
These questions are integrated mathematically in the following equation:

$$Score_{i,j,t} = \frac{n.worse\ value_{j,t} + \frac{n.\ same\ value_{i,j,t}}{2}}{n.with\ a\ value_t} \tag{1}$$

This equation calculates the score by dividing the number of companies in a certain year for a certain subcategory with a lower score, plus half of the companies with the same score, by the total number of companies with an assigned score for that subcategory in a certain year. In this equation $score_{i,j,t}$ stands for the score for company *i*, for subcategory *j*. Furthermore, *n.* stands for the number of companies. All ESG scores have are on a scale of 0 to 100, however due to some observations being left out of the data and some scores having more values of 0, the average of the scores can differ between the subcategories. To compute the overall combined ESG score, the scores for the different subcategories are assigned a category weight. The reason for this is that the subcategories are computed based on a number of indicators that relate to a specific subcategory. The number of indicators can vary between the subcategories. In order to understand the scoring methodology, all the different subcategories of ESG scores and their labels in the data are summarized below. The weights and the number of indicators corresponding to the subcategories of ESG scores are reported in Table 3.1. A total of 178 indicators are used to compute the scores, however number of indicators can differ per subcategory. This causes some subcategories to be more important for determining the combined ESG score than others. As a result, the number of indicators attributed to a subcategory correlates with its given weight. The weight is calculated by dividing the number of indicators used to compute a certain score, by the total number of indicators, as can be seen in equation 2.

$$Weight_i = \frac{Indicators_i}{Total\ number\ of\ indicators} \tag{2}$$

In this equation $Weight_i$ stands for the weight for subcategory *i*, and $Indicators_i$ stands for the number of indicators used to determine the score for subcategory *i*.

Table 3.1. Number of indicators and the corresponding weights for the three main ESG pillars and their subcomponents

Pillar	Subcategory	Number of Indicators in Scoring	Weights (%)
Environmental	Resource Use	20	11
	Emissions	22	12
	Innovation	19	11
Social	Workforce	29	16
	Human Rights	8	4.5
	Community	14	8
	Product Responsibility	12	7
Governance	Management	34	19
	Shareholders	12	7
	CSR Strategy	8	4.5
Total		178	100

Note. Source: Thomson Reuters, 2018.

There are a number of different ESG scores reported in the data and they have different corresponding variable names. The variable named esg stands for the overall ESG score and it is computed using all scores for the subcategories. There is also a combined ESG score which is comparable to the overall ESG score, except it assigns weights to the subcategories in order to compute the overall ESG score. Table 3.1 lists the different subcategories for the three main pillars of ESG which are used to compute the overall ESG scores (esg), and the combined ESG scores (esgc). The environment pillar consists of the following sub scores: resource use score (esgru), emissions score (esge), and environmental innovation score (esgi). The social pillar consists of the following scores: workforce score (esgw), human rights score (esghr), community score (esgcom), and product responsibility score (esgpr). And lastly the pillar for governance contains the following sub scores: management score (esgm), shareholders score (esgs), and CSR strategy score (esgcsrs).

3.4 Defining the Main Variables

Khan (2022) summarizes a number of common variables used in ESG research. In the paper the variables return on assets (ROA) and return on equity (ROE) are used as proxies to measures for a firm's financial performance. This study replicates the variable for firm performance by using these same measures for firm performance, which is calculated by dividing the earnings before interest, tax, depreciation and amortization (EBITDA), by the total assets, as shown in equation 3. Figure 3.1 plots the ROA and the ROE in separate histograms on the left. It shows that these variables are subject to large outliers. To account for this, the two variables are transformed into a natural logarithm, which makes the variables more normally distributed. The results of these transformations can be seen in the two scatterplots on the right side of the figure, next to the scatterplots in which the return on assets and return on equity are shown. Equation 3 shows the calculation for the ROA.

$$\ln(ROA_{i,t}) = \ln\left(\frac{EBITDA_{i,t}}{Total\ Assets_{i,t}}\right) \tag{3}$$

This equation calculates the firm performance measured by the return on assets, using EBITDA for company *i*, in year *t*. In addition, the natural logarithm of this ratio is taken to deal with a large number of outliers and skewness, which are caused by large differences in company sizes and large differences in the earnings of the companies. The ratio for firm performance accounts for a few differences between companies, in order to make it easier to compare them. For example, all monetary values in the Compustat database are reported in the company's local currency. By making it a ratio, the differences in exchange rates between countries are accounted for.

$$\ln(ROE_{i,t}) = \ln\left(\frac{EBITDA_{i,t}}{Total\ Common\ Ordinary\ Equity_{i,t}}\right) \tag{4}$$

The ROE calculated in equation 4 is almost identical to the ROA as seen in equation 3, except that it divides the EBITDA by the total common ordinary equity instead of the total assets for firm *i* in year *t*. The ROE is another proxy for firm performance.

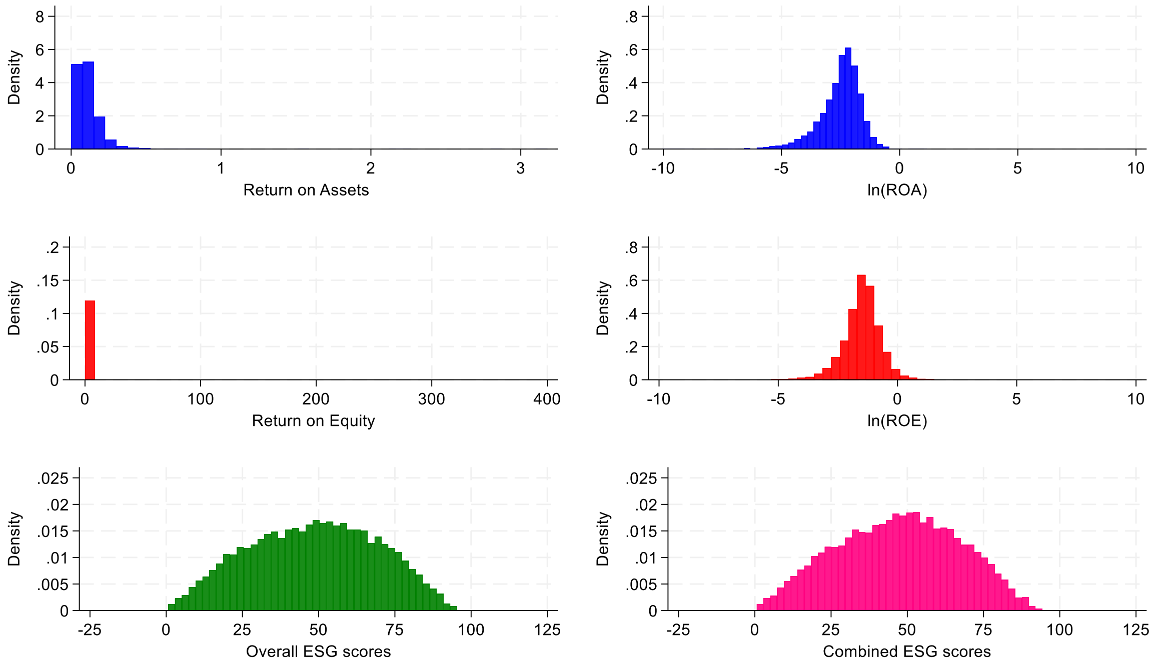

Figure 3.1. Histograms showing the distribution of observations for the main variables of interest: Return on Assets (ROA), Return on Equity (ROE), and the Overall ESG scores on the left side, and the natural logarithm of ROA and ROE, and the Combined ESG scores on the right

Figure 3.1 shows six histograms with density plots. Top left: return on assets and its density. Top right: distribution of the natural log-transformed return on assets. Middle left: return on equity. Middle right: log-transformed return on equity. Bottom left: histogram of the overall ESG scores, distributed on a scale of 0 to 100. Bottom right: combined ESG scores, also on a scale of 0 to 100. The observations for the years 2018 to 2021 are pooled together in the histograms.

Khan (2022) also mentioned a few variables to differentiate companies based on firm characteristics. For example the variable leverage. This variable will be measured as the debt to equity ratio, which is calculated by taking the total liabilities and dividing this by the total common ordinary equity for company *i*, in year *t*, which is visualized by equation 5.

$$\ln(Leverage_{i,t}) = \ln\left(\frac{Total\ Liabilities_{i,t}}{Total\ Common\ Ordinary\ Equity_{i,t}}\right) \tag{5}$$

Due to the many outliers in the data after calculating the leverage ratio, this variable was also transformed into a natural logarithm. A reason to include the debt to equity ratio can be that companies differ in their capital structure. One of the ESG scores is shareholders score. Companies with more equity compared to debt can be more concerned with the needs of their shareholders. These companies can also experience more pressure from their shareholders to guide their company policies to improve their ESG scores.

3.5 Other Variable Definitions

To control for other differences between firms, a number of additional control variables are included in the data. These include: size of the company, country of the headquarters of the company, the year in which the data was measured and reported, and the industry in which the firms mainly operate. The variable size is measured as total revenue, however due to a few large outliers, this variable is also transformed into a natural logarithm. This variable can take on only positive values. Industry is a categorical variable that uses a string of 6 numbers to identify the different industries. The industries are classified based on the Global Industry Classification (GIC) standard. It consists of 11 sectors, 25 industry groups, 74 industries, and 163 subindustries. The code originally consists of 8 numbers, if the sub-industry is also included. The first 2 digits stand for the sector, the second pair of 2 digits stands for the industry group, and the third pair of 2 digits stands for the industry. Lastly the 2 digits at the end of the 8 digit code stand for the subindustry. The variable industry can account for fundamental differences between industries. For example, some industries can be more capital intensive than others. This leads to larger differences in the return on assets, due to a larger value of total assets compared to the EBITDA. The industry that a firm is operating in can also determine which subcategories of ESG scores are more important. An IT company that is mostly operating on the digital landscape may be more focused on the Social and Governance pillars, instead of the Environmental pillar of ESG, while an oil company's score is probably more dependent on the Environmental pillar for their overall ESG score.

The year in which the ESG scores are computed can also account for possible differences, because the ESG scores are based on a percentile rank scoring methodology (Thomson Reuters, 2017). This makes the ESG scores for a certain company within a certain year partially depend on the performance of other companies within that same year, not only the firm's own ESG performance. The factors that can play a role in these possible differences can be derived from equation 1. The total number of companies that are used to calculate the percentile rank score can differ between the years for which the ESG scores are reported in the dataset. This can lead to different outcomes for the same company between multiple years, even if their performance had been the same, throughout those same years.

The last variable, which is the country of the headquarters of the company. This variable is also a categorical variable, which bases their country codes on the International Organization for Standardization (ISO)'s 3166 country code standard. This variable links the headquarters of a company to the country it is located in. This variable can account for differences in the rules and regulations that the companies have to adhere to, which are dependent on the country in which the headquarters is located. Section 3.6 reports the summary statistics for some of the variables defined above.

3.6 Summary Statistics

This section shows the summary statistics and the correlations of the variables that will be included in this study, to get a brief understanding of the data. Table 3.2 shows the summary statistics, Table 3.3 shows the correlations between the different variables, and Table 3.4 shows the correlations between the variables for the different subcategories of the ESG measures.

Table 3.2. Summary statistics of the main variables of interest

Variable	Obs.	Mean	Std. dev.	Min.	Max.
ln(ROA)	18,792	-2.534	0.897	-10.050	1.156
ln(ROE)	18,792	-1.505	0.862	-9.255	5.862
ln(leverage)	18,792	0.291	1.262	-8.647	7.233
ln(firm size)	18,788	9.262	2.967	0.732	19.449
esg	18,792	48.348	20.810	0.600	95.430
esgc	18,792	47.095	19.838	0.600	94.270
esgru	18,792	46.503	31.814	0	99.940
esge	18,792	48.908	31.500	0	99.940
esgi	18,792	28.584	31.814	0	99.890
esgm	18,792	52.301	28.096	0.100	99.940
esgs	18,792	51.848	28.106	0.050	99.960
esgcsrs	18,792	45.608	31.244	0	99.940
esgw	18,792	59.876	27.361	0.140	99.930
esghr	18,792	37.270	33.657	0	99.300
esgcom	18,792	46.559	30.396	0	99.940
esgpr	18,792	51.107	31.172	0	99.960

Note. All variables starting with esg represent ESG scores. The ESG scores can take on a value of 0 to 100. Some minimum and maximum ESG scores are not 0 and 100. This is caused by the exclusion of some observations. The abbreviation obs. refers to the number of observations, std. dev. Stands for standard deviation. Min. and max. stand for the minimum value and the maximum value.

Table 3.2 shows the summary statistics of the non-categorical variables included in the data. It is noticeable that the mean, minimum, and maximum values for the subcategories of ESG scores differ from each other. This is caused by the exclusion of some observations due to missing values and the multiple years of data being combined within the summary statistics. Another noticeable value in the summary statistics is the mean for the variable esgi_s. It is much lower with a value of 28.577, compared to the other scores. The other scores have a mean ranging between 46 and 53. The low value for the mean is due to many scores with a value of zero given to companies for this subcategory. The scores that don't have a value of zero are uniformly distributed between a score of 0 to 100. Another noticeable value is the mean for the variable esgw_s. It's mean is 59.836, which seems to deviate from the mean. It is expected based on equation 1 that the average of the scores should be close to 50, however the sample that is included in this dataset for esgw_s seems to be performing better on average than the population. This is probably caused by excluding observations that had missing values for the Compustat data, which also had low values for the scores for this subcategory.

Table 3.3. Correlations between the firm performance measures, the overall and combined ESG scores, and the control variables

Variable	ln(ROA)	ln(ROE)	esg	esgc	ln(lev.)	ln(f. size)
ln(ROA)	1.000					
ln(ROE)	0.652	1.000				
esg	0.000	0.155	1.000			
esgc	0.012	0.145	0.963	1.000		
ln(lev.)	-0.388	0.385	0.217	0.192	1.000	
ln(f. size)	-0.037	0.119	0.257	0.232	0.214	1.000

Note. 18,788 observations are used to calculate the correlations. The variable ln(lev.) stands for ln(leverage) and ln(f. size) stands for ln(firm size). Esg refers to the overall ESG scores, and esgc refers to the combined ESG scores. The Table reports the correlations matrix between the different variables reported in the data, which are used for this study.

Table 3.3 reports the correlations between the measures for firm performance: return on assets (ROA) and return on equity (ROE), the overall ESG scores, the combined ESG scores, and the variables leverage ratio and firm size. The correlation between firm performance and the overall ESG scores seems to be nonexistent with a correlation of zero. Leverage however seems to have a positive correlation of 0.217 with the overall ESG scores. The correlation between the overall ESG scores and the combined ESG scores are very strong, with a value of 0.963, which seems reasonable because the overall and combined ESG scores are calculated using the same measures for ESG.

Table 3.4. Correlations between the ESG score subcategories

Variable	esgru	esge	esgi	esgm	esgs	esgcsrs	esgw	esghr	esgcom	esgpr
esgru	1.000									
esge	0.783	1.000								
esgi	0.454	0.446	1.000							
esgm	0.310	0.296	0.181	1.000						
esgs	0.111	0.095	0.062	0.197	1.000					
esgcsrs	0.652	0.639	0.374	0.334	0.105	1.000				
esgw	0.697	0.690	0.347	0.332	0.122	0.608	1.000			
esghr	0.662	0.594	0.363	0.291	0.120	0.509	0.579	1.000		
esgcom	0.564	0.510	0.315	0.316	0.121	0.470	0.551	0.579	1.000	
esgpr	0.525	0.501	0.338	0.205	0.076	0.403	0.510	0.426	0.413	1.000

Note. This table shows the correlations between each subcategory of the ESG measures, which are used to calculate the overall and combined ESG scores. 18,792 observations were used to calculate the correlations. The table reports the correlations matrix between the different subcomponents of ESG scores.

Table 3.4 reports the correlations for the subcategories of ESG scores. The correlations between the subcategories and firm performance are very low, with the highest positive correlation being 0.059 for esghr, and the highest negative correlation being -1.103 for esgi. Debt to equity has higher correlations compared to firm performance, although the strongest correlation is 0.197 for esgru and none of those correlations are negative. The correlations between the scores for the subcategories of the ESG pillars fluctuate strongly between 0.1 and 0.8. This suggest that some scores are more strongly associated to one another than other scores. A possible explanation for the differences in correlations between the scores for the subcategories can be found in the differentiation between the three main pillars of ESG.

Figure 3.2 shows the overall and combined ESG scores plotted against the measures of firm performance: ln(ROA) and ln(ROE). Each dot in the scatterplots represents an observation in the data. It seems that the observations for the overall and combined ESG scores are distributed evenly across the different values for ln(ROA) and ln(ROE), which suggests a very low correlation between ESG scores and firm performance. The visual representation of the data in these scatterplots are in line with the correlations presented in Table 3.3.

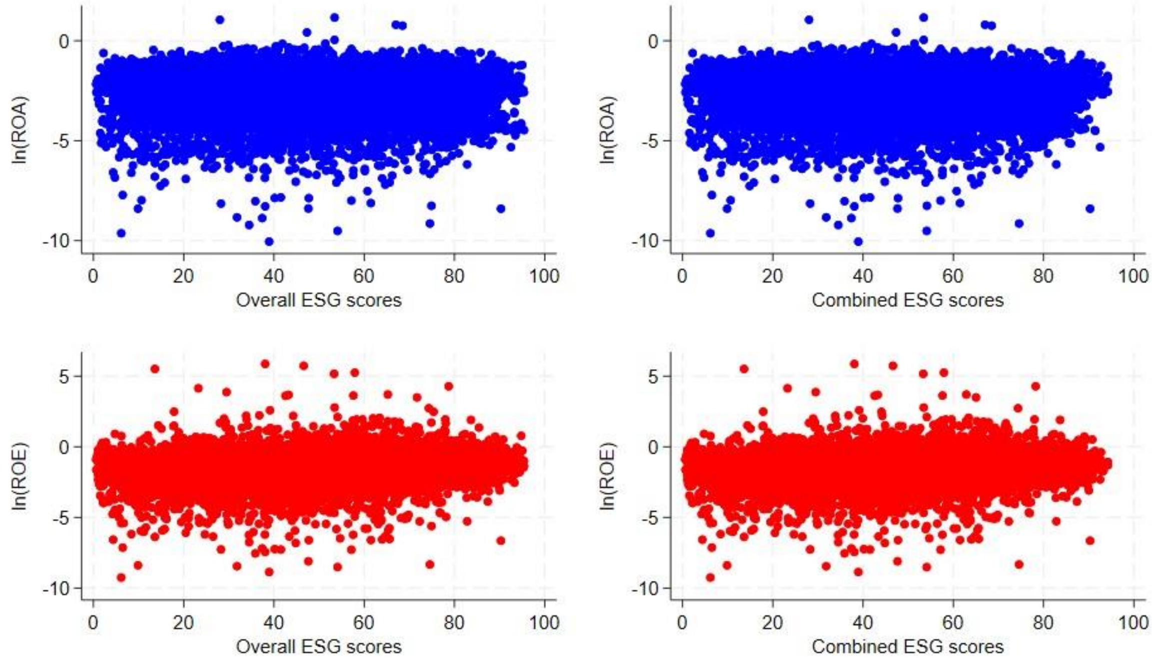

Figure 3.2. The overall and combined ESG scores plotted against the measures for firms' financial performance, ln(ROA) and ln(ROE)

Figure 3.2 shows four scatter plots. Top left: scatter plot with overall ESG scores on the X-axis and the values for the natural log of return on assets on the Y-axis. Top right: scatter plot with combined ESG scores on the X-axis and the natural log of return on assets on the Y-axis. Bottom left: scatter plot with overall ESG scores on the X-axis and the natural log of return on equity on the Y-axis. Bottom right: scatter plot with combined ESG scores on the X-axis and the natural log of return on equity on the Y-axis. Each dot represents an observations in the data. The years 2018 to 2021 are pooled together in the scatter plots. The scores on the X-axis are on a scale from 0 to 100.

CHAPTER 4 Method

4.1 Panel Data Analysis

The question if higher ESG scores are associated with a better financial firm performance, is going to be tested using three different hypotheses: Higher overall ESG scores have significant positive effects on the firm's financial performance; higher combined ESG scores have a significant positive effect on a firm's financial performance; and lastly one or more of the ESG scores for the subcategories have a significant effect on the firm's financial performance. This study uses panel data, which consists of cross-sectional and time-series data. This means that multiple firms are observed over multiple time periods. There are three main models suitable for the analysis of panel data, the pooled regression, the fixed effects model, and the random effects model. These three methods all use the Ordinary Least Squares (OLS) method. The pooled regression combines all observations across time and across firms into one single cross-sectional dataset, rather than accounting for a panel data structure. This method of analysis assumes that there are no differences between the firms on the basis of their characteristics, and it also assumes that there are no differences between time periods. In the case of this study, it is highly unlikely that there are no such differences in characteristics between firms. A formal F-test can be performed, which tests the null hypothesis that all the individual firm unobserved fixed effects are zero. If this null hypothesis is rejected, it means that there is evidence that one or more of these individual fixed effects are not zero. It suggests that there are differences between the firms. This justifies including fixed effects in the model to account for these differences between the firms.

4.2 Fixed Effects

The fixed effects model controls for unobserved heterogeneity by including a set of dummy variables, also called fixed effects, for each individual firm in the data (Brooks, 2019). These fixed effects capture all characteristics of the firms that do not vary over time, but may affect the dependent variable. The fixed effects model is formulated mathematically as follows:

$$y_{it} = \alpha_i + \beta x_{it} + u_{it} \quad (6)$$

In this equation y_{it} is the dependent variable for firm *i* at time *t*. α_i is the firm specific fixed effect for firm *i*. x_{it} is the vector of time-varying explanatory variables. β is the vector of coefficients to be estimated and u_{it} is the error term, which varies across entities and time periods. The error term could be written as:

$$u_{it} = \mu_i + v_{it} \quad (7)$$

Where μ_i is the firm specific effect, and v_{it} is the remaining disturbance term, that varies over time and between firms, which is left unexplained.

4.3 Random Effects

The random effects model proposes differences in the intercepts between the firms, and these intercepts are again assumed to be constant over time. The difference between the fixed effects model and the random effects model is that the random effects model assumes that each cross-sectional unit is assumed to arise from a common intercept and a random variable. This random variable varies cross-sectionally, but is still constant over time.

$$y_{it} = \alpha + \beta x_{it} + \omega_{it} \qquad (8)$$

In this equation y_{it} is the dependent variable for firm i at time t. α is the common intercept for all firms. x_{it} is the vector of time-varying explanatory variables. β is the vector of coefficients to be estimated and ω_{it} is the error term, which varies across entities and time periods. The error term could be written as follows:

$$\omega_{it} = \epsilon_i + v_{it} \qquad (9)$$

Where ϵ_i is the firm specific effect, and v_{it} is the remaining disturbance term, that varies over time and between firms, which is left unexplained. There are no dummy variables which capture the heterogeneity in the cross-section. In the random effects model, this is captured by the term ϵ_i. The random effects model has one major drawback. In order for it to be valid, the error term ω_{it} has to be uncorrelated with all the explanatory variables. This means that all the x_{it} should be uncorrelated with the ϵ_i and v_{it}. In other words, the unobserved omitted variables need to be uncorrelated with the included explanatory variables. If this assumption is violated, the parameter estimated will be biased and inconsistent.

4.4 Hausman Test

To determine if the fixed or random effects model is most suitable for this particular study, a version of the Hausman test can be performed. It tests whether the unobserved individual effects are correlated with the regressors in the model. If the null hypothesis of the Hausman test is rejected, then there is evidence of correlation between the individual effects and the regressors. In this situation, the fixed effects model is more appropriate. If the null hypothesis of the Hausman test is not rejected, there is not enough evidence of correlation between the regressors and the error term, which makes the random effects model more appropriate. To perform the Hausman test, both the fixed and random effects models need to be estimated and compared to one another.

4.5 Model Specifications

To study the effect of the ESG scores on a firm's financial performance and to test the hypotheses stated in section 4.1, fixed effects models will be estimated, similar to the study performed by Nguyen et al. (2022). When testing if the fixed or random effects were more important, the results indicated that fixed effects were more appropriate. The results and implications for these tests are discussed in more detail in the results section of the paper. The following fixed effects models are going to be estimated:

$$\ln(ROA_{i,t}) = \alpha_i + \beta_1 ESG\ overall\ score_{i,t} + \beta_2 \textbf{Control Variables} + \mu_i + v_{i,t} \quad (10)$$

$$\ln(ROA_{i,t}) = \alpha_i + \beta_1 ESG\ combined\ score_{i,t} + \beta_2 \textbf{Control Variables} + \mu_i + v_{i,t} \quad (11)$$

$$\ln(ROA_{i,t}) = \alpha_i + \beta_1 ESG\ subcategory\ scores_{i,t} + \beta_2 \textbf{Control Variables} + \mu_i + v_{i,t} \quad (12)$$

$$\ln(ROE_{i,t}) = \alpha_i + \beta_1 ESG\ overall\ score_{i,t} + \beta_2 \textbf{Control Variables} + \mu_i + v_{i,t} \quad (13)$$

$$\ln(ROE_{i,t}) = \alpha_i + \beta_1 ESG\ combined\ score_{i,t} + \beta_2 \textbf{Control Variables} + \mu_i + v_{i,t} \quad (14)$$

$$\ln(ROE_{i,t}) = \alpha_i + \beta_1 ESG\ subcategory\ scores_{i,t} + \beta_2 \textbf{Control Variables} + \mu_i + v_{i,t} \quad (15)$$

In these equations $\ln(ROA_{i,t})$ is the dependent variable in the first three models, which is estimated for firm *i* at time *t*. The last three models use $\ln(ROE_{i,t})$ as the dependent variable, which is another measure for firm performance. The variables $\beta_1 ESG\ overall\ score_{i,t}$, $\beta_1 ESG\ combined\ score_{i,t}$, and $\beta_1 ESG\ subcategory\ scores_{i,t}$ are the independent variables which are of main interest in explaining firm performance. The term $ESG\ subcategory\ scores_{i,t}$ consists of all the individual scores for the subcategories that make up the combined score, and β_1 represents all the corresponding coefficients. The control variables are included in the model as the term $\beta_2 \textbf{Control Variables}$. This term consists of the variables $\ln(Debt\ to\ Equity_{i,t})$ and $\ln(Firm\ Size_{i,t})$, and β_2 represents coefficients that are associated with these control variables. The term for the fixed effects is α_i and it stands for the individual firm characteristics that are time invariant, but affect the dependent variable. The term for the fixed effects is u_i and it represents the firm specific error term. It stands for the unexplained variation of the model for firm *i* and it includes the unobserved factors that impact the dependent variable, but are constant over time. Lastly there is the error term $v_{i,t}$. This term stands for the remaining unexplained variation that is firm specific and varies over time. Note that in these equations, the categorical variables for industry and country are excluded. The reason for this is multicollinearity, which will arise when these variables and the individual fixed effects are both included in the model for each firm. The term α_i already controls for these variables by including the fixed effects, meaning that each firm has its own intercept which is estimated using a dummy variable. This intercept is assumed to be constant over time, which can also be assumed for the categorical variables industry and country.

The hypotheses that will be tested using the fixed effects regression equations are the following: Higher ESG scores have significant positive effects on the firm's financial performance. In this case the ESG scores will be measured using the overall ESG scores and the combined ESG scores. The combined ESG scores are calculated using weights, hence the difference between the overall and combined scores. This will be tested using equations 10, 11, 13, and 14. Lastly equations 12 and 15 test the following hypothesis: one or more of the ESG scores for the subcategories have a significant effect on the firm's financial performance. The last hypothesis will be tested by including all the different scores for the subcomponents of ESG. These are included in the equations as subcategory scores, which represent the collection of individual scores. The effects on firm performance will be tested using two measures, ROA and ROE. Equations 10 to 12 will test the hypotheses using ROA as dependent variable and equations 13 to 15 will test the same hypotheses using ROE as dependent variable. Two additional hypotheses will be tested which state the following: there are significant differences in the impact of higher ESG scores on firms' financial performance between sectors; and larger firms have a greater impact of higher ESG scores on their financial performance compared to smaller firms. These will be tested by including an interaction term between the different sectors and the ESG scores, and between firm size and the ESG scores.

To determine which kind of effects, random or fixed, is more appropriate to use when estimating the models, models with a similar structure as equations 10 to 15 will be estimated in order to perform a Hausman test. These models have the same structure as the models presented by equation 10 to 15, however they will include the categorical variables country and industry, and random effects instead of the fixed effects. Different variations of the models will additionally be estimated, for example with lagged values of the dependent variables to test the robustness of the results. A disadvantage of including the lagged values is that a fourth of the observations will be lost, because the data consists of only 4 years. This may lead to less accurate estimates. Interaction terms will also be included in some of the models, for which variants will be estimated in order to again perform a Hausman test, to determine which kind of effects should be included. The next section presents the estimated results for these models.

CHAPTER 5 Results

The relationship between ESG scores and firm performance, is studied using the 6 models described in the methodology section. They are estimated and their results are reported in the Tables 5.1 and 5.4. Variations of these models are estimated to check the robustness of the estimates and are reported in Tables 5.2, 5.3, and A.1 from the Appendix. Table 5.1 reports the three models that use ROA as dependent variable. Table 5.2 excludes the control variables to see if this changes the estimated results for the coefficients of the independent variables. Table 5.3 reports the same models as Table 5.1, but with lagged values instead of the current values for the independent variables. Table 5.4 reports the same models, but instead of using ROA, it uses ROE as dependent variable to measure firm performance. Lastly Table 5.5 reports 4 additional models that include interaction terms. The estimated models in this study include fixed and random effects models, which are estimated using the OLS method. These fixed and random effects models are then compared to one another using the Hausman test and the more appropriate model is reported in the results. The dependent variable in all the models is transformed into a natural logarithm. This means that a one-unit increase in the independent variable, in this case the ESG score measures, is associated with a percentage change in the dependent variable, in this case the measures for firm performance. The control variables for firm size and the debt to equity ratio are also transformed into a natural logarithm. When they change by 1%, the dependent variable changes by the percentage change of the coefficient for these control variables. The constant term in these equations is the fixed effect, which differs for each firm.

5.1 Return on Assets

Table 5.1 presents the results for the fixed effects regression models that use return on assets as dependent variable. The table reports 3 models, the first includes the overall ESG score as the independent variable, the second one includes the combined ESG score, and the third one includes the ESG scores for all the subcategories. There are 18,638 observations used to estimate the three models. The within R2 is reported in the table for the three models below the number of observations. It measures the proportion of the variance in the dependent variable that is explained by the model within each firm over time. Model 1 and 3 have a value of 0.172 and the value for Model 2 is 0.171 for the within R2. This indicates that the model is bad at explaining the variation in return on assets within a firm. The P-value for the Hausman test is also reported for each model. A significant Hausman test statistic for a P-value that is smaller than 0.05 indicates that there is evidence of correlation between the individual effects and the regressors. This is the case for all three models in the table, because they all have a P-value of 0.000, which is lower than 0.05. It can be concluded that the fixed effects model is more appropriate to estimate for these models in this study.

Table 5.1. OLS regression results for the models with ROA as dependent variable and including fixed effects

Variable	Dependent variable		
	ln(ROA)	ln(ROA)	ln(ROA)
	(1)	(2)	(3)
esg	0.250		
	(0.471)		
esgc		0.196	
		(0.429)	
esgru			0.164
			(0.408)
esge			-0.571
			(0.360)
esgi			-0.470
			(0.313)
esgm			0.195
			(0.261)
esgs			-0.424
			(0.259)
esgcsrs			0.495^{*}
			(0.291)
esgw			-0.028
			(0.400)
esghr			-0.035
			(0.303)
esgcom			0.039
			(0.329)
esgpr			0.482^{*}
			(0.286)
ln(leverage)	-0.235^{***}	-0.235^{***}	-0.235^{***}
	(0.011)	(0.011)	(0.011)
ln(firm size)	0.657^{***}	0.657^{***}	0.658^{***}
	(0.014)	(0.014)	(0.014)
Fixed Effects	Yes	Yes	Yes
Number of Observations	18,638	18,638	18,638
Within R2	0.172	0.171	0.172
Hausman P-value	0.000	0.000	0.000

Note. Standard errors are reported between parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The independent variables for ESG scores, the variables starting with esg, can take on a value between 0 and 100. The coefficient estimates and standard errors for the variables esg, esgc, esgru, esge, esgi, esgm, esgs, esgcsrs, esgw, esghr, esgcom, and esgpr are scaled up by 1000.

Model 1 in Table 5.1 reports an insignificant coefficient of 0.00025 for esg, which stands for the overall ESG score. This suggests that when the overall ESG score increases with one unit, the return on assets changes with 0.00025% in the sample. The standard error for this coefficient is 0.000471 and its P-value is larger than 0.1. The estimates for esg suggests that the return on assets is not significantly affected by a change in the overall ESG score for the firms in the sample. The insignificant P-value of the coefficient shows that there is not enough evidence to conclude that the coefficient is different from 0. The same holds true for the variable esgc in Model 2, which measures the combined ESG score, and is calculated using the individual scores for the subcategories of ESG measures and their corresponding weights as reported in Table 3.1. Model 3 from Table 5.1 shows a significance of a P-value that is larger than 0.1, meaning a 10% significance level, for the independent variables esgcsrs and esgpr. This is still not enough evidence to conclude that they have a significant effect on return on asset. The other ESG measures in Model 3 are not significant at all. This suggest that there is no single individual ESG measure that in itself has a significant effect on return on assets.

The control variables ln(leverage) and ln(firm size) are significant at the 1% level. The leverage ratio has a significant negative coefficient of -0.235 in relation to return on assets in all three models, and firm size has a significant positive coefficient with a value of 0.657 in Model 1 and Model 2 and 0.658 in Model 3. This suggests that an increase of 1% in leverage is associated with a decrease of 0.235% of ROA in the sample. A 1% increase in firm size is associated with a 0.657% increase in ROA in the sample. When controlling for leverage and firm size, and including fixed effects, the overall, combined, and individual ESG scores are not found to have a significant relationship with the firms' financial performance, which is measured as ROA.

5.2 Excluding Control Variables

Table 5.2 shows the same regression models as Table 5.1, however in these models the control variables are excluded. The within R2 measure has worsened with a maximum value of 0.001 for Model 3. This indicates that this model is even worse at explaining the variation in the return on assets compared to the models that include the control variables leverage and firm size. Such a low R2 suggests that these models have almost the same predictive power as a model that has only a constant term and thus consists of only a straight line that fits the data. The Hausman test statistic has become insignificant for all three models in Table 5.2, indicating that random effects are more appropriate for a model with these specific variables.

The significance of the independent variables in the three models has not changed. This shows that in these models the ESG scores still show no evidence of having a significant effect on return on assets, when leverage and firm size are not controlled for. When comparing the coefficients for the ESG scores from Table 5.2 to the ones from Table 5.1, it can be seen that the scores have changed signs, when relevant control variables are added, however they are still insignificant. This provides more evidence to suggests that there is no significant relationship between ESG scores and the performance of firms in this sample, for the 4 years for which the firms were recorded in the data.

Table 5.2. OLS regression results for models with ROA as dependent variable, including fixed effects, and excluding control variables

		Dependent variable	
Variable	ln(ROA)	ln(ROA)	ln(ROA)
	(1)	(2)	(3)
esg	-0.093		
	(0.517)		
esgc		-0.096	
		(0.471)	
esgru			0.169
			(0.449)
esge			-0.288
			(0.396)
esgi			-0.458
			(0.343)
esgm			0.103
			(0.287)
esgs			-0.298
			(0.284)
esgcsrs			0.458
			(0.319)
esgw			-0.433
			(0.439)
esghr			-0.053
			(0.333)
esgcom			-0.062
			(0.362)
esgpr			0.316
			(0.314)
Fixed Effects	Yes	Yes	Yes
Number of observations	18,642	18,642	18,642
Within R2	0.000	0.000	0.001
Hausman P-value	0.858	0.838	0.741

Note. Standard errors are reported between parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The independent variables for ESG scores, the variables starting with esg, can take on a value between 0 and 100. The coefficient estimates and standard errors for the variables esg, esgc, esgru, esge, esgi, esgm, esgs, esgcsrs, esgw, esghr, esgcom, and esgpr are scaled up by 1000.

5.3 Lagged Values

Table 5.3 presents the same models as Table 5.1, but in these models the lagged values of the independent variables are used instead of the current values. The reasoning is that it may take some time for the effects of the ESG scores of the past year to start having an effect on the firm performance, and thus the effects may start to be visible in the data the year after. The values in these models are lagged by one period, which corresponds to one year. This causes the models to lose a quarter of all observations. The total number that is used to estimate these models is 12,143 observations. This may lead to less accurate estimations for the coefficients.

The within R2 is higher for these models than the models without lagged values reported in Table 5.1. The value for the three models in Table 5.3 for the within R2 is 0.215, compared to 0.171 and 0.172 for the models in Table 5.1. This suggests that the models with lagged values have more predictive power than the models without lagged values, however it is only a small difference. The Hausman value for the models in Table 5.3 are all highly significant. This means that the fixed effects are still appropriate for these models compared to random effects. The coefficients for all of the ESG score variables are still insignificant in these models, which still gives no evidence in support of the hypothesis that higher ESG scores constitute to better firm performance.

Table 5.3. OLS regression results with ROA as dependent variable, including fixed effects and lagged values for the independent variables

		Dependent variable	
Variable	ln(ROA) (1)	ln(ROA) (2)	ln(ROA) (3)
L.esg	0.499		
	(0.570)		
L.esgc		0.292	
		(0.516)	
L.esgru			-0.696
			(0.502)
L.esge			0.440
			(0.441)
L.esgi			-0.008
			(0.378)
L.esgm			0.540*
			(0.321)
L.esgs			0.178
			(0.314)
L.esgcsrs			0.221
			(0.353)
L.esgw			0.182
			(0.488)
L.esghr			-0.015
			(0.368)
L.esgcom			-0.237
			(0.401)
L.esgpr			0.016
			(0.347)
ln(leverage)	-0.305***	-0.305***	-0.306***
	(0.017)	(0.017)	(0.017)
ln(firm size)	0.746***	0.746***	0.746***
	(0.018)	(0.018)	(0.018)
Fixed Effects	Yes	Yes	Yes
Number of Observations	12,143	12,143	12,143
Within R2	0.215	0.215	0.215
Hausman P-value	0.000	0.000	0.000

Note. Standard errors are reported between parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The independent variables for ESG scores, the variables starting with esg, can take on a value between 0 and 100. The coefficient estimates and standard errors for the variables esg, esgc, esgru, esge, esgi, esgm, esgs, esgcsrs, esgw, esghr, esgcom, and esgpr are scaled up by 1000.

5.4 Return on Equity

Table 5.4 presents the same models as Table 5.1, but instead of using return on assets as measure for firm performance, the models in this table use return on equity as measure for firm performance. The number of observations to estimate these models is 18,638. This is the same as the models in Table 5.1. The within R2s for the models in Table 5.4 are little higher with the values 0.198 and 0.199 compared to the R2s of 0.171 and 0.172 for the models in Table 5.1. This suggests that the variables are a little better in explaining the variation in return on equity compared to return on assets. The Hausman test is still highly significant, making fixed effects appropriate for these models. An interesting part is the control variables. In these models, they all have a positive sign. According to these models a 1% increase in the leverage ratio leads to a 0.314% increase in return on equity in data sample. A 1% increase in firm size leads to a 0.63% increase in return on equity. All the coefficients for the ESG measures are again insignificant at the 5% confidence level. Using a different measure for firm performance gives the same results as the other models, making the results more robust. From these results it cannot be concluded that better ESG scores have a significant positive effect on firm performance. This is due to the lack of significant results.

A number of extra robustness checks are done to test the robustness of the results. Table A.1 from the appendix shows the same models as Table 5.1, but robust standard errors are added to control for heterogeneity in the errors. The results from Table A.1 don’t show any large differences with Table 5.1, with regards to the standard errors, the signs of the coefficients, and the significance of the coefficients. This implies that the model’s underlying assumptions about the standard errors are reasonable met. In other words, the standard errors that are estimated seem to be valid. However, it is difficult to know this for certain, because it may still be the case that not all assumptions about the standard errors do in fact hold. Thus the results should be interpreted with care.

Table 5.4. OLS regression results with ROE as dependent variable and including fixed effects

	Dependent variable		
Variable	ln(ROE) (1)	ln(ROE) (2)	ln(ROE) (3)
esg	0.376		
	(0.475)		
esgc		0.305	
		(0.433)	
esgru			0.217
			(0.412)
esge			-0.578
			(0.364)
esgi			-0.480
			(0.316)
esgm			0.235
			(0.264)
esgs			-0.503*
			(0.261)
esgcsrs			0.345
			(0.294)
esgw			0.012
			(0.403)
esghr			0.083
			(0.306)
esgcom			0.125
			(0.332)
esgpr			0.499*
			(0.289)
ln(leverage)	0.314***	0.314***	0.314***
	(0.011)	(0.011)	(0.011)
ln(firm size)	0.630***	0.630***	0.630***
	(0.014)	(0.014)	(0.014)
Fixed Effects	Yes	Yes	Yes
Number of Observations	18,638	18,638	18,638
Within R2	0.198	0.198	0.199
Hausman P-value	0.000	0.000	0.000

Note. Standard errors are reported between parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The independent variables for ESG scores, the variables starting with esg, can take on a value between 0 and 100. The coefficient estimates and standard errors for the variables esg, esgc, esgru, esge, esgi, esgm, esgs, esgcsrs, esgw, esghr, esgcom, and esgpr are scaled up by 1000.

5.5 Interaction Effects

Lastly Table 5.5 presents the regression results with models that have included interaction terms between the ESG scores and different sectors. These models test if there are differences in the effects of ESG scores between sectors, as indicated by Mashayekhi et al. (2024). The many interactions between the industries would be very chaotic when all put into one regression model. So instead of including all the different industries, the industries are grouped together in different sectors, which are indicated by the identification numbers 15 to 60. The interaction terms show how the relationship between the ESG scores and the firms' financial performance can differ bewteen sectors. The coefficient of -0.009 for esg in the Model 1 represents the within-entitiy effect, in this case within a particular firm, over time and suggests that, on average, a one-unit increase in the ESG score is associated with a 0.9% decrease in ROA, when holding all other variables constant. In Model 2 for the variable esgc this corresponds to a 0.7% decrease in the sample. Both these coefficients are significant at the 1% confidence level. Sector 10 is used as a reference category and is therefore omitted form the model. It refers to the Energy Sector.

The interaction terms can be interpreted as follows, for sector 15, when ESG score increases with one unit, the return on assets increases with -0.009 + 0.010 = 0.001, or 0.1% in the sample. Or it increases with 1% more compared to sector 10. All coefficients for the interaction terms are significant at the 5% confidence level except for sector 25 in Model 2. Some sectors seem to indicate a positive relationship between higher ESG scores and firm performance for example sector 30, with a coefficient of 0.002 when adding it to the independent variable. Other industries that show a positive significant coefficient estimate are the sectors: 40, 50, 55, and 60. Other industries seem to indicate a negative relationship in both models, for example sector 10 and sector 25. The rest of the sectors are neutral, with the coefficient being neutral, positive, and or negative in the two models. These results indicate that there are differences in the effect of ESG scores on a firm's financial performance between industries within the sample. For some industries a higher ESG score seems to indicate a higher return on assets compared to the energy sector, in other industries higher ESG scores seem to be related to lower return on assets compared to the energy sector.

Model 3 and 4 in Table 5.5 include an interaction term between ln(firm size) and the ESG scores. The coefficient for the independent variable for ESG score is again insignificant. The coefficient for the interaction term could be interpreted as follows. When the firm size increases with 1%, the positive effect of ESG score on ROA decreases by 0.020%. The constant for this model would in this case be unrealistic because it is highly unlikely that the firm size would be 0. In Model 3 and 4 the coefficient for the interaction term is insignificant, thus it cannot be concluded that differences in firm size significantly impact the effects of ESG score on the firms' financial performance.

Table 5.5. OLS regression results with ROA as dependent variable, interaction terms, and fixed effects

	Dependent variable			
Variable	ln(ROA)	ln(ROA)	ln(ROA)	ln(ROA)
	(1)	(2)	(3)	(4)
esg	-0.009***		0.471	-0.010***
	(0.003)		(1.546)	(0.004)
esgc		-0.007***		
		(0.003)		
sector#esg				
15. Materials	0.010***	0.007**		0.010***
	(0.004)	(0.003)		(0.004)
20. Industrials	0.009***	0.007**		0.009***
	(0.004)	(0.003)		(0.004)
25. Consumer Discretionary	0.008**	0.006*		0,008***
	(0.004)	(0.003)		(0.004)
30. Consumer Staples	0.011***	0.009***		0.011***
	(0.004)	(0.003)		(0.004)
35. Health Care	0.008**	0.006**		0.008***
	(0.004)	(0.003)		(0.004)
40. Financials	0.011***	0.008***		0.011***
	(0.004)	(0.003)		(0.004)
45. Information Technology	0.009***	0.007**		0.009***
	(0.004)	(0.003)		(0.004)
50. Communication Services	0.012***	0.010***		0.012***
	(0.004)	(0.003)		(0.004)
55. Utilities	0.011***	0.009***		0.011***
	(0.004)	(0.003)		(0.004)
60. Real Estate	0.012***	0.009***		0.012***
	(0.004)	(0.003)		(0.004)
ln(leverage)	-0.235***	-0.235***	-0.219***	-0.235***
	(0.011)	(0.011)	(0.011)	(0.011)
ln(firm size)	0.656***	0.656***	0.695***	0.654***
	(0.014)	(0.014)	(0.015)	(0.016)
ln(firm size)#esg			-0.020	0.044
			(0.161)	(0.165)
constant	-8.569***	-8.569***	-8.922***	-8.552***
	(0.130)	(0.129)	(0.138)	(0.145)
Fixed Effects	Yes	Yes	Yes	Yes
Number of Observations	18,638	18,638	18,788	18,638
Within R2	0.173	0.173	0.194	0,173
Hausman P-value	0.000	0.000	0.000	0.000

Note. Standard errors are reported between parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The independent variables for ESG scores, the variables starting with esg, can take on a value between 0 and 100. In Model 1 the variable esg is interacted with the categorical variable sector, and in Model 2 the variable esgc is interacted with the variable sector. Sector 10 represents the Energy sector and is used as reference category. The coefficients for the ESG score variables: esg, and ln(firm size)#esg in Model 3 and 4 are scaled up by 1000.

CHAPTER 6 Discussion

The trend in the interest in ESG investing that Abhayawansa & Tyagi (2021) have noticed, can be seen in the availability of data for this study. There has been an increase in ESG data that is being reported around the start of the COVID-19 pandemic. For this reason the years 2018 to 2021 are used in this study. For these years this study does not find enough evidence for the hypothesis that higher ESG scores increase firms' financial performance within firms. This finding is not in line with Friede et al. (2015) and Nguyen et al. (2022), who did find a positive relationship using the same measures for firm performance, which are ROA and ROE. The study performed in this paper used a similar method as Nguyen et al. who also performed a least squares regression with fixed effects. This is contrary to Friede et al. who performed a meta-analysis. The results from the study performed in this paper is more in line with the results found by Halid et al. (2023). They reviewed multiple studies and found different results, ranging from studies reporting positive relationships, neutral relationships, and negative relationships. It shows that there is not an overall consensus on the effectiveness of ESG scores and that the results of different studies could vary. Billio, et al. (2021) discuss this point in their paper. They show that there is heterogeneity in ratings criteria for ESG. This could be an explanation for the differences in the results between studies. The results may depend on which ESG framework is used to determine the scores and which rating agency has given the scores, because two agencies can give two different ESG scores to the same company according to Billio, et al..

Mashayekhi et al. (2024) have looked at the individual ESG subcomponents and their results show that the social and economic pillars are the two fundamental pillars of ESG performance in all industries in general. However this can differ from industry to industry. The study performed in this paper also looked at the individual subcomponents of ESG, but did not find enough evidence to conclude that any subcomponent of ESG has a significant positive or negative relationship with the firms' financial performance within a firm. However when including interaction terms between the different sectors that firms can operate in and the overall ESG scores, the results showed that there are significant differences in the effects of ESG scores on the firms' financial performance. The sign however differs per industry and could be positive as well as negative, depending on the sector, when using the energy sector as reference category. This is in line with the observation of Mashayekhi et al. that the effects of ESG subcomponents can differ per industry. Some sub measures may be more appropriate for certain industries, and certain industry-specific characteristics could also lead to fundamental differences in the average ROA and ROE between industries.

Lastly an interaction term between firm size and ESG scores was added to the models to test if there are any differences in the effect of ESG scores on firms' financial performance for companies with different sizes. The coefficients for the interaction terms between firm size and ESG scores were however insignificant. From these results it cannot be concluded that there are significant differences of the impact of ESG scores on firms' financial performance for firms with different sizes. This finding does not support the results of Gregory (2022) who found a uniform positive relationship between firm size and ESG over ratings agencies in his study.

CHAPTER 7 Conclusion

This paper has studied the relationship between ESG scores and their effect on firms' financial performance, measured as return on assets and return on equity. There has been an extensive literature that studies this relationship with different papers having different outcomes. Some report a positive relationship, while others report a negative or neutral relationship. The study performed in this paper differentiates the ESG scores into overall ESG scores, combined ESG scores, and subcomponents of ESG scores that make up the overall and combined ESG scores. The data consisted of panel data for the years 2018 to 2021 and an OLS panel regression with fixed effects was used to analyze the data. All the estimated models included fixed effects.

The results did not show any significant relationship between higher ESG scores and better firm performance within firms over time. When differentiating between the subcomponents of ESG scores, there was not a single significant coefficient found at the 5% significance level. This indicates that in the timespan of the sample for this study, there is not enough evidence to conclude that better ESG scores lead to better financial firm performance within firms, when considering the entire sample. However, when including interaction terms between the ESG scores and the different sectors that firms operate in, the ESG score coefficients were all found to be significant except for one. This suggests that the effects that higher ESG scores have on a firm's financial performance differs per sector, hence explaining the insignificant results in the models that do not include the interaction terms. To summarize this study does not find any significant effect of higher ESG scores on the firms' financial performance within firms when considering the full sample, however it does indicate that there are significant differences in the relationship between ESG scores and firms' financial performance between sectors when interaction terms are included in the models.

These findings could help firms and investors understand the implications and effectiveness of ESG scores when used as a measure to evaluate companies. This study suggests that when approaching ESG scores, it would not be appropriate to use ESG scores as a global measure, but instead sector-specific characteristics should be take into account when evaluating the impact of ESG scores on a firm's financial performance. These results should however be interpreted carefully, because there could be limitations with the data. For example, selection bias could be present when selecting the firms to include in the sample. Firms with bad ESG scores could choose not to report ESG scores, which automatically excludes them from the data. Another problem that could arise is omitted variable bias. Some relevant variables may not be observed or not included in the models, which could lead to endogeneity issues. This means that the estimates may be biased and inconsistent, which could lead to wrong interpretations for the significance of the coefficients. This makes it uncertain that this study has measured a causal relationship. Lastly the short duration of the sample may be an explanation for the mostly insignificant results, because it may take a few years for the effects of the ESG policies within companies to take effect and be measurable in the data.

Future research could look at the effects of ESG on firm performance as measured by Tobin'sQ. This way the firm performance could be measured on a market value bases, which takes into account the stock value of a company relative to the total assets. It essentially compares the market value to the book value of a company. This measure includes the expectations of shareholders and can give a forward looking indication of the ESG scores on the firms' financial performance and prospects, instead of only a performance that is solely based on historical data. In addition, a longer time duration for the sample may show different results. The implementation of ESG policies may require more time, and it may also take more time for the ESG policies to be impactful in improving the firms' ESG scores, and in turn their financial performance.

REFERENCES

Abhayawansa, S., & Tyagi, S. (2021). Sustainable Investing: The Black Box of Environmental, Social and Governance (ESG) ratings. *Social Science Research Network*. https://doi.org/10.2139/ssrn.3777674

Amel-Zadeh, A., & Serafeim, G. (2018). Why and How Investors Use ESG Information: Evidence from a Global Survey. *Financial Analysts Journal, 74*(3), 87–103. https://doi.org/10.2469/faj.v74.n3.2

Barnett, M. L., & Salomon, R. (2006). Beyond dichotomy: the curvilinear relationship between social responsibility and financial performance. *Strategic Management Journal, 27*(11), 1101–1122. https://doi.org/10.1002/smj.557

Billio, M., Costola, M., Hristova, I., Latino, C., & Pelizzon, L. (2021). Inside the ESG ratings: (Dis)agreement and performance. *Corporate Social-responsibility and Environmental Management, 28*(5), 1426–1445. https://doi.org/10.1002/csr.2177

Bloomberg Intelligence. (2023). ESG assets may hit $53 trillion by 2025, a third of global AUM.

Brooks, C. (2019). *Introductory Econometrics for Finance.* https://doi.org/10.1017/9781108524872

Chen, S., Song, Y., & Gao, P. (2023). Environmental, social, and governance (ESG) performance and financial outcomes: Analyzing the impact of ESG on financial performance. *Journal of Environmental Management*, 345, 118829. https://doi.org/10.1016/j.jenvman.2023.118829

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. https://doi.org/10.1080/20430795.2015.1118917

Giese, G., Lee, L., Melas, D., Nagy, Z. T., & Nishikawa, L. (2019). Foundations of ESG Investing: How ESG affects equity valuation, risk, and performance. *Journal of Portfolio Management, 45*(5), 69–83. https://doi.org/10.3905/jpm.2019.45.5.069

Gregory, R. P. (2022). The influence of firm size on ESG score controlling for ratings agency and industrial sector. *Journal of Sustainable Finance & Investment, 14*(1), 86–99. https://doi.org/10.1080/20430795.2022.2069079

Halid, S., Rahman, R. A., Mahmud, R., Mansor, N., & Wahab, R.A. (2023). A Literature Review on ESG Score and Its Impact on Firm Performance. *International Journal of Academic Research in Acounting Finance and Management Sciences, 13*(1), 272-282. https://doi.org/10.6007/ijarafms/v13-i1/15101

Khan, M. A. (2022). ESG disclosure and Firm performance: *A bibliometric and meta-analysis. Research in International Business and Finance, 61*, 101668. https://doi.org/10.1016/j.ribaf.2022.101668

Khan, M., Serafeim, G., & Yoon, A. (2016). Corporate Sustainability: First evidence on materiality. *The Accounting Review*, *91*(6), 1697–1724. https://doi.org/10.2308/accr-51383

Mashayekhi, B., Asiaei, K., Rezaee, Z., Jahangard, A., Samavat, M., & Homayoun, S. (2024). The relative importance of ESG pillars: A two-step machine learning and analytical framework. *Sustainable Development*. https://doi.org/10.1002/sd.2979

Nguyen, D. T., Hoang, T. G., & Tran, H. G. (2022). Help or Hurt? The Impact of ESG on Firm Performance in S&P 500 Non-Financial Firms. *Australasian Accounting Business and Finance Journal, 16(*2), 91–102. https://doi.org/10.14453/aabfj.v16i2.7

Thomson Reuters, & (2018). Thomson Reuters ESG scores. https://www.esade.edu/itemsweb/biblioteca/bbdd/inbbdd/archivos/Thomson_Reuters_ESG_Scores.pdf

Velte, P. (2017). Does ESG performance have an impact on financial performance? Evidence from Germany. *Journal of Global Responsibility*, *8*(2), 169–178. https://doi.org/10.1108/jgr-11-2016-0029

APPENDIX A Additional Robustness Checks

Table A.1. OLS regression results with ROA as dependent variable, the models include robust standard errors and fixed effects

		Dependent variable	
Variable	ln(ROA)	ln(ROA)	ln(ROA)
	(1)	(2)	(3)
esg	0.250		
	(0.480)		
esgc		0.196	
		(0.411)	
esgru			0.164
			(0.402)
esge			-0.571
			(0.383)
esgi			-0.470
			(0.315)
esgm			0.195
			(0.279)
esgs			-0.424*
			(0.246)
esgcsrs			0.495
			(0.306)
esgw			-0.028
			(0.439)
esghr			-0.035
			(0.338)
esgcom			0.039
			(0.314)
esgpr			0.482*
			(0.284)
ln(leverage)	-0.235***	-0.235***	-0.235***
	(0.025)	(0.025)	(0.025)
ln(firm size)	0.657***	0.657***	0.658***
	(0.037)	(0.037)	(0.037)
Fixed effects	Yes	Yes	Yes
Number of observations	18,638	18,638	18,638
Within R2	0.172	0.171	0.172

Note. Standard errors in are reported between parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The independent variables for ESG scores, the variables starting with esg, can take on a value between 0 and 100. The coefficient estimates and standard errors for the variables esg, esgc, esgru, esge, esgi, esgm, esgs, esgcsrs, esgw, esghr, esgcom, and esgpr are scaled up by 1000.

Table A.1 shows the regression results that have been estimated using robust standard errors. The significance of the results don't differ from the results reported in Table 5.1. This shows that including robust standard errors does not significantly change the results. The results give more support for the assumption that the standard errors are robust and valid. However there may still be other assumptions that can be violated, which can cause the standard errors of the estimated coefficients to be biased, leading to wrong interpretations of the significance of the coefficients.