ERASMUS UNIVERSITY ROTTERDAM
ERASMUS SCHOOL OF ECONOMICS
Bachelor Thesis Economics & Business

Foundation for ESG Investing? A Panel Data Analysis of Western Europen Firms' ESG Ratings and Market Performance

Author: Roko Medanic
Student number: 562236
Thesis supervisor: dr. Fabrizio Core
Second reader: dr. Ruben de Bliek
Finish date: 01.08.2023

The views stated in this thesis are those of the author and not necessarily those of the supervisor, second reader, Erasmus School of Economics or Erasmus University Rotterdam.

ABSTRACT

Do ESG ratings fundamentally affect a firm's equity performance in Western Europe? To answer the question as concretely as possible a decade worth of data (2013-2022) on firm ESG ratings, returns, and specific control variables was collected from Eikon Refinitiv. The comprehensive sample entailed 10,695 observations, spread across 1,972 different firms, located in 14 different Western European countries. Employing multiple regression models, firm returns were in essence regressed on ESG rating variables to ascertain the significance, sign, and magnitude of their influence. The findings revealed the following key results: (1) On average, combined ESG ratings, Social ratings, and Environmental ratings exhibit a negative association with stock returns in Western Europe, while Governance ratings have no discernible effect on returns. (2) The magnitude of the ESG variables' relationships is notably minimal though, potentially negligible depending on the reader's perspective. (3) Regionally, subsamples of Western European countries do not register statistically significant effects for the ESG rating variables, and the ones that do are nearly exclusively negative, indicating that for the majority of regions share prices are unresponsive to ESG ratings. (4) Moreover, it appears that the relationship between ESG ratings and firm returns is primarily driven by endogenous factors. These findings have important implications for firms, investors, and policy makers. Our results caution against an overly enthusiastic and uncritical adoption of using ESG ratings as a determinant for making investment decisions. A prevalent notion in much literature suggesting that ESG focused portfolios outperform markets, or that highly ESG-rated firms outperform lowly ESG-rated ones, may be potentially misleading. Accordingly, stakeholders should exercise prudence and bear in mind a more comprehensive approach in their decision-making processes, accounting for multifaceted factors that drive equity performance in Western Europe.

Keywords: Environmental, Social, and Governance (ESG) Scores & Performance, Western Europe**,** Panel Data

JEL codes: G11, G32, G34, G38, M14, Q51

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 Introduction

By 2020 88% of publicly listed companies had ESG initiatives in place (James d'Ath, 2022), and by 2022 nearly all top MBA programs offered ESG courses (Yao, 2022). It is evident ESG metrics serve a significant solution within the business world to address the crucial sustainability challenges our generation is currently confronted with. Formally, ESG metrics are non-financial firm performance metrics that measure a firm's environmental, social, and governance performance. In combination with other financial indicators, they allow investors, analysts, and other stakeholders to evaluate a company's sustainability and long-term performance holistically, gaining a more complete picture of its overall value and risks. As a result, understanding the relationship between ESG metrics and firm financial performance is key for all stakeholders.

Whilst this seems like a straightforward relationship, one of the largest problems the field of research is faced with are contradicting findings. For example, Whelan et al. (2021) conducted a meta-analysis of over 1000 research papers published in the United States (US) between 2015 and 2020, but only found "a positive relationship between ESG and financial performance for 58% of the 'corporate' studies focused on operational metrics such as ROE, ROA, or stock price" (Whelan et al., 2021, p.2). De facto this raises the question, how can ESGs form such an integral parts of business curriculums and firm policies but fail to establish a relationship between a firm's ESG metrics and its financial performance for over 40% of studies?

Numerous avenues to examine exist, yet economic research tends to focus on US markets (Gavrilakis and Floros, 2023), obviously driven by the high concentration of global research funds at higher educational institutions in the US. However, considering indications that ESGs may have a larger impact on European markets, such as that Europe invests over 20% more of its total assets into sustainable and responsible investment strategies compared to the US (Global Sustainable Investing Review, 2016), examining Western European markets over US markets becomes a fundamental principle of this paper. Specifically, this paper tackles the relationship between firm ESG scores and stock returns in Western Europe. On a basic level, stock returns, because unlike other financial metrics such as Profits, ROI, and ROA, they entirely capture the public's opinion. Ultimately, it is uncertain whether firms face a trade off between financial performance and sustainability, or whether sustainability aids financial performance. Hence, the object of this study is firm ESG metrics and their effect on firm stock returns in Western Europe.

Understandably, in recent history the relationship between firm sustainability metrics and stock performance has been a hotly researched field given the importance of a transition to more sustainable

practices by firms and in general, society. Derwall, Guenster, Bauer, and Koedijk (2005) analyse if environmental responsibility influences a company's share returns using Innovest ratings. They analyse US firm stock market returns from 1997 to 2003 using the Carhart (1997) four-factor model, and through a high-low approach for the Innovest ratings show a considerable outperformance of high-rated firms over low-rated firms.

Halbritter et al. (2015) publish a similar paper, also based on US markets, investigating "the link between corporate, social, and financial performance and the environmental, social, and corporate governance (ESG) ratings", and additionally review "existing empirical evidence pertaining to this relationship" (Halbritter et al., 2015, p.1). The authors also utilized the Carhart (1997) four-factor model, however, unlike earlier research presented by Koedijk (2005), a significant difference in the returns between high ESG rated firms and low ESG rated firms was not established. Their research further extended others' by applying Fama and MacBeth (1973) regressions to micro analyse the ESG pillars. The regressions revealed a significant positive influence of several ESG variables, but the authors concluded that investors are hardly able to exploit this relationship.

Findings in the US generally seem to be contradictory. Given we know that Western European countries consistently rank superior to the US in terms of sustainability, it is prompt to question whether research finds similar inconclusive results there. Lucia et al. (2020) examine a Western Europen data sample, exploring the predictive ability of ESG scores on a firm's ROA and ROE using a combination of machine learning and inferential models. Whilst their firm-level financial performance indicator is not stock market returns, their findings are that machine learning accurately predicts the ROA and ROE of firms in Western Europe, indicating the existence of a positive relationship between ESG practices and financial indicators.

Even in Western Europe however, literature is contradictory. Gavrilakis and Floros (2023) examined the effects of ESG scores on the stock performance of large capped firms within Western Europe. Similarly, to our paper, one of their motivations to analyze a European sample was the inherent lack of European market-based studies. They did not take a portfolio or meta-analysis approach but rather enlisted Ordinary Least Square (OLS) regressions. Specifically, their regressions included a total ESG score parameter as well as market capitalization, price-to-book value, and Sharpe ratio parameters. Their results illustrated that investing in high ESG scoring firms does not affect stock returns, apart from in Italy where the effect is negative.

Evidently, literature also lacks definitive conclusions on Europe. Nevertheless, the relative scarcity of research on Western Europe coupled with the notable sustainability disparities between Europe and the US, justifies producing a novel scientific study which contributes to the pool of

European-based market research. Specifically, this paper chooses to build on the methodology entailed by Gavrilakis and Floros (2023). Their simplistic methodology of selecting basic regression analyses seems most appropriate given the contradictory nature of currently available literature. It does not seem logical to analyze complex models and intricate details of a relationship if a fundamental relationship is inherently non-existent. The study by Gavrilakis and Floros (2023) was constrained by a few issues, however. Two explicitly apparent deficiencies worth immediately mentioning are (1) their sample representativeness and (2) not micro-analyzing the effects of the respective ESG pillars. Critically, their sample was unrepresentative of Western Europe as they selected firms from six European countries of which three are considered Economically troubled and two are not necessarily recognized as leaders of sustainability: Portugal, Italy, Greece, Spain, France, and Germany.

As an ancillary consideration, readers may raise the inquiry pertaining to the justification for prioritizing the examination of stock performance over alternative financial indicators. The rationale is founded on the sheer degree of direct influence investors wield over companies' share prices. Considering the well-established significant relationship between markets and media channels and recognizing the considerable prominence of sustainability positive topics and agendas in the media and public consciousness over the last three decades, it is reasonable to presume that the effects of ESG scores on firm financial performance may be more explicitly highlighted in the fluctuations of share prices than, revenue, for instance.

The research question follows: ***Does a fundamental relationship exist between a firm's ESG scores and a firm's stock market returns?*** As stated, to examin this relationship, I will enrol similar regression analyses used by Gavrilakis and Floros (2023), extending their methodology and literature by: (1) Studying the relatively unresearched European firms, with a dataset which is more representative of Western Europe (London Stock Exchange (LSE), Euronext (ENXB), Deutsche Borse (DB1), SIX Swiss Exchange (SIX), NASDAQ Nordic (NDAQ), and Baltic Exchange (OMX). (2) Micro-analysing ESG scores' effects on a pillar level to define a more detailed relationship between ESG rating and firm returns. (3) Providing robustness checks for endogeneity issues and subsample variation (which entails micro-analyzing the relationship on a country level as different geographical regions may pronounce relationships to different degrees). And (4) improving the model by enlisting panel techniques. The combination of these characteristics will provide novel up-to-date findings relevant to literature.

Importantly, as highlighted and discussed by Gavrilakis and Floros (2023), the standard control variables used will be market capitalizations (controlling for the size of firms is important as larger firms may exhibit different stock return patters compared to smaller firms), price-to-book values (controlling for differences in valuation levels among companies, accounting for the possibility that the market may

have priced in the impact of ESG metrics on stock returns based on their valuation), and Sharpe ratios (controlling for difference in risk profile among companies).

In terms of data, everything was collected from Eikon Refinitiv. This is due to the fact that its database claims to be 99.8% accurate and to cover over 80% of the global market cap for more than 450 different ESG measures (Breitz and Partapuoli, 2020). Moreover, Gavrilakis and Floros (2023) exclusively used Eikon Refinitiv data which makes our study more comparable to theirs. Specifically, I extracted yearly data for the following variables: firm ESG scores (including specific E-S-G pillar scores and combined ESG Scores), stock returns, market capitalizations, price-to-book values, and Sharpe ratios. The data was collected across the last decade, year-end, (2013-2022) for firms from the following exchanges: LSE, ENXB, DB1, SIX, OMX and NDAQ. Briefly, methodologically Refinitiv calculates the total ESG scores as a weighted sum of the individually calculated pillar scores, which are calculated through 10 further specified subcategories, of which is grounded on a further subset of 186 metrics. The use of exchanges and data gives us reliable sample that is a holistic representation of Western European countries.

Overall, this paper endeavors to investigate the potential influence of firm ESG scores on their stock performance within the context of Western Europe. As a basis, a firm's ESG ratings inform investors of how sustainable a firm is, which can be informative to an investor on multiple levels. For example, considering current trends towards more sustainable practices, especially in Europe where regulators are stricter, being unsustainable can potentially harm a firm's future expected growth and stock performance. To an extent, ESG scores can help quantify those types of concerns for investors. Hence, on a rudimentary level, the ***primary hypothesis*** is that a positive relationship between ESG ratings and firm stock returns will be found.

Furthermore, I postulate a ***second hypothesis*** that posits the environmental pillar score will exert a more positive influence over stock returns compared to the combined ESG score and remaining social and governance pillar scores, respectively. This conjecture stems from the fact that currently the largest global threat to humanity is climate change. Consequently, countries have rapidly changing policies and quotas regarding firm emissions and environmental impacts. Hence, when evaluating a firm's sustainability, the environmental impact of a firm should be of greater concern than the social and governance ratings. Thus, ceteris paribus, comparatively, a firm's environmental rating should be a stronger indicator of less risky future cash flows, resulting in improved equity performance. Furthermore, I believe that the same deductions should apply to the Environmental score and combined ESG score, namely the weights in Refinitv's methodology used to calculate the combined ESG score to do not favor environmental scores.

Additionally, and finally, I ***thirdly hypothesize*** that the effects of ESG scores on firm stock performance will be more pronounced for Nordic countries. This is a conclusion I base on the fact that Nordic countries consistently rank highest in sustainability rankings, hence ESG scores may play a more significant role in the core values of their business models, consequently affecting their stock performance.

In this research I show that there does not seem to be a relationship present between ESG ratings and returns for firms. But if one were critical, it would be possible to state that there is a minute negative relationship which may be difficult to practically exploit. Additionally, regional specifications also follow these arguments, with insignificant coefficients dominating. In general, it seems that if a positive relationship were to be found it would be endogenously driven and possibly subject to data mining. As such, various stakeholders should be wary when interpreting ESG ratings within the context of an investment decision.

The remainder of the paper is structured as follows. Firstly, Section 2 presents a detailed explanation of the relationship examined and its constituents, so that it is comprehendible to all readers. Then, section 3 presents the data, followed by section 4 which describes the methodology enlisted to analyze that data. Next, section 6 illustrates our results, and finally section 7 concludes our research demonstrating the inferences drawn from our results followed by its implications and limitations.

CHAPTER 2 Theoretical Framework

2.1 Stock Market Returns

The introduction may cover certain terminology and topics that may not be familiar to everyone, hence, as a starting point of this theoretical framework it is important to clarify the two main concepts that form our relationship of interest. The first is stock market returns, and though trivial, being the outcome variable that this study aims to measure, it is important to define. Classically, "a stock market return is the positive or negative change in value of an investment or asset over time. A positive return means a profit has been made on the investment. A negative return means that there has been a loss on the investment. Total stock market returns include dividends and interest payments, as well as the price change in the stock. The change in stock price alone is referred to as a nominal return" (MoneySense.com, 2023). From now on this paper will refer to the stock market returns of a firm as simply its *returns*.

Importantly to note, is that whilst the definition above refers to returns in the sense of the profitability of a sole investment, this paper will treat it as a financial firm performance metric for public firms. This is due to the fact that a firm's stock price is the direct valuation of the firm as perceived by external investors, constituting the likes of institutional traders and retail investors. Categorically, returns fall under the umbrella term of "financial" firm performance metrics because it is one of multiple ways to measure how well a firm is doing, e.g., revenues, profits, operating margins, and debt-to-equity ratios (Sage.com, 2023).

Academic literature is more practical, and returns are typically not explicitly defined, but rather are understood to be mathematically derived as the current closing price divided by the previous closing price, minus 1 (yahoo!finance, 2023). This outputs the return as a percentage, dependent only upon the time frame between the numerator and denominator and whether the author opts to use adjusted closing prices to correct for dividends and other firm decisions that may change the closing price of their share (as this paper will do).

Furthermore, given that returns will be on the LHS of our regressions, it is important to understand what can lead to changes in the returns of a firm, specifically, what leads to changes in the share price of a firm. To tackle this, it is most practical to take a mathematical approach and examine the two most famous models for valuing a firm and its share price: (1) Williams (1938), with the Discounted Cash Flow model of valuing a company and therefore indirectly calculating its share price, and (2) Gordon (1959), with the Gordon Growth Model of directly calculating a firm's share price. The

formulas for both models can be found under Appendix A.1 and A.2. Upon inspection two common terms exist across both models, the discount rate (r) and growth rate (g); mathematically it is evident that both majorly influence the underlying stock price of a firm. Specifically, a higher growth rate and a lower discount rate both independently increase the stock price.

Thus, to understand the latter part of the theoretical framework it is vital to grasp what the discount rate and growth rate of a firm represent in the sense of its share price. The Corporate Finance Institute (CFI) defines the discount rate as "the rate of return used to discount future cash flows back to their present value. This rate is often a company's Weighted Average Cost of Capital (WACC), required rate of return, or the hurdle rate that investors expect to earn relative to the risk of the investment." In essence it is a factor that allows investments to be standardized in terms of perceived riskiness, so that one may compare different investment options (CFI.com, 2023). As such, factors deemed to make an investment riskier will result in a heavier discount rate, and so will provide a smaller return, or in our case a lower share price. On the flip side, the growth rate is defined as "the constant rate at which a firm's expected free cash flows are assumed to grow indefinitely" (CFI.com, 2023). Therefore, a firm with higher prospects is likely to have a share price that matches that. Which effect dominates the share price is a widely discussed topic though, Zimmermann et al. (2021) find that both the implied growth rate and discount rate have large and significant effects on stock returns but that ultimately the effect of the growth rate of a firm's cash flows dominates.

Intuitively it would be considerate to discuss what exactly influences the discount rate and growth rate, however, the discussion would be endless given the broad nature of the topic. Consequently, it is more appropriate to use one's intuition in understanding whether certain factors would make firms more risky or positively prospective from an investment standpoint, or, precisely, their ability to generate future cash flows.

2.2 ESG Metrics

The second variable in our relationship of interest is our predictors of interest, the ESG's (environmental, social, and governance). Historically, they came to light in the early 1960s when investors simply began excluding stocks or entire industries from their portfolios based on business activities such as tobacco production or involvement in the South African apartheid regime, labelling it then as socially responsible investing. It was a move towards investing into more ethically sustainable companies. Since then, sustainability has grown in popularity, with the first mention of ESGs in a popular report titled "Who Cares Wins" in 2004 where a joint initiative of some of the 20 largest companies in the financial industry presented recommendations "to better integrate environmental,

social and governance issues in analysis, asset management and securities brokerage" (The Global Impact, 2004). Formally, the ESG framework was only created in 2010 and later introduced in 2015. But that has not stopped most professional databases from collecting and offering ESG data from earlier years, 2002 to date, though.

Today, Gartner- management consulting company- defines ESG metrics (our predictor variable) as "a collection of corporate performance evaluation criteria that assess the robustness of a company's governance mechanisms and its ability to effectively manage its environmental and social impacts. Examples of ESG data include the quantification of a company's carbon emissions, water consumption or customer privacy breaches" (Gartner.com, 2023). Academic definitions in literature are identical given the nature of a universal metric.

ESG scores tend to come from large market data providers who all generally adopt a similar framework for the calculations of their various ESG scores and sub-scores. For example, Refinitiv's methodology is designed to transparently and objectively measure companies' relative ESG performance, commitment, and effectiveness. And, as briefly mentioned in the introduction, the scores are based on publicly reported data, incorporating over 630 company-level ESG measures, of which a subset of 186 comparable and material measures per industry are selected, and then grouped down to 10 categories contributing to the three E-S-G pillar scores, and ultimately the total ESG score. One element present in generally all ESG metric calculations is the presence of weights which vary for each of the categories and pillars across each industry.

Akin to how our outcome variable is a firm performance metric, so is our predictor. In contrast however, it is a non-financial performance metric, which in the past may have led some to deem it less relevant since it is not a direct numerical monetary representation of a firm's wellbeing. However, with the ever-growing relevancy of sustainability, non-financial metrics such as ESG scores are integral to firm analysis and wellbeing.

2.3 Tying ESG Metrics to Market Returns

What motivates the relationship between ESG scores and returns, and what has that relationship looked like in the past? Whilst substantial literature on our topic of interest exists, the variables may not always be identical to ours. There exist many similar sustainability indices and corporate performance metrics that help quantify the sustainability of a firm's practices and its financial health. Important, is that the fundamental relationship discussed, highlighting the underlying mechanisms driving the relationship, are coherent and logical.

Homoeconomicus rationally prefer to take on little risk, maximizing their risk adjusted return. Firms with higher ESG scores exhibit sustainable characteristics and should consequently be perceived as less risky resulting in easier access to capital (ceteris paribus), appreciating the firm's share price. As touched upon towards the end of chapter 2.2, KPMG established that 68% of banks incorporated ESG criteria in its lending policies today, concluding that a more sustainable firm or project is more likely to obtain a lower cost of debt (KPMG Property Lending Barometer, 2022). Similarily, Giererd-Potin et al. (2013) illustrated that firms with better socially responsible (SR) scores had lower costs of equity as "investors ask for an additional risk premium when they accept to hold non-socially responsible stocks" (Giererd-Potin et al., 2013, p.1). It is by little surprise then that Wong et al. (2021) found the ESG certification of firms in Malaysia is negatively associated with cost of capital. This illustrates the investors' enjoyment of sustainable investments due to the perceived reliable cash flows. Deriving the share price of a firm from the standard equations presented by both the Gordon Growth model and Discounted Cash Flow model, in Appendix A.1 and A.2, a lower cost of equity or capital causes an increase in the share price of a firm (ceteris paribus). Hence, a firm who's ESG score grows should also witness a growth in its share price, therefore, positive returns year on year with respect to ESG scores. As such, it is vital to isolate the effects of risk from ESG ratings when examining firm returns.

Interrelatedly, to achieve a high sustainability index score, a firm must truly be sustainable. Meaning, companies prioritizing their sustainability (ESG factors) must adopt more efficient and innovative practices to achieve this. For example, implementing eco- or energy-friendly technologies as well as waste reduction measures. Being sustainable then, by default, should lead to cost savings and increased operational efficiency, increasing profitability and thus stock performance (ceteris paribus). Aroul et al. (2022) examined the relationship between ESG performance of Real Estate Investment Trusts funds (REITs) and their operational efficiency and performance, running a standard ordinary least square (OLS) estimation on all publicly traded REITs on the US exchange between 2019-20. The analysis indicated that REITS scoring higher on the ESG scale have higher operational efficiency and performance. Additionally, the association between ESG scores and operational performance is stronger for firms which have higher operational efficiency. As such, companies recording higher ESG scores minimize costs associated to any activities on an operational level. Facets like machinery and employee productivity are maximized in terms of their productivity whilst maintaining a level of quality or timeliness. In essence, there is less waste of energy translating into a cost-effective firm and a more streamlined business model. Therefore, increasing their operational performance and, ceteris paribus, their share price.

Investor sentiment and market perception are also two intertwined concepts which link ESG scores to returns. Ford et al.'s (2022) examined the potential link between ESG factors and option

market-derived short-term investor sentiment. Using portfolio sorts and dynamic panel regressions they tested whether ESG factors are related to short-term investor sentiment for US companies. Most importantly, their results revealed that the difference between the highest and lowest ESG rated portfolio was negative, meaning that highly rated ESG firms received significantly more optimistic sentiment than the lowly ESG rated firms. By deduction then, ESG scores play a role in influencing investor sentiment and market perception through signaling to investors a firm's ability to manage environmental and social challenges effectively. This leads to a favorable perception within the investment community, increasing demand for such firms and appreciating its share price. Ford et al.'s (2022) additionally found that among the three ESG pillars only the environmental score was a significant factor for option traders', with a higher environmental score significantly improving sentiment. Critically, the authors concluded that the most sophisticated option traders seriously consider the environmental risk associated with individual companies when investing. Consequently, our analysis may also find a stronger positive association between the environmental pillar score of firms and their returns. Theoretically, this may be explained by the adamant focus on environmental sustainability by governments and the media over the last 30 years, due to climate change, giving rise to the importance of the environmental score over the remaining two pillars to investors.

Moreover, among other factors, there is a differentiation between countries in terms of culture, mindset, and general society. This heterogeneity should be reflected in variation of influence of ESG scores on investor preferences, sentiment, and market perception among nations. Consequently, the influence of these factors on firm share prices through ESG scores is expected to vary as well. To what extent is unknown given the broad nature of societal and cultural impact on decision making, however previous studies like Gavrilakis and Floros (2023) indeed found variation in the effects of ESG scores on returns for different countries with changes in sign and magnitude present.

The link between ESG scores and returns is indirect. That is, there are variables influencing ESG scores which are correlated with the error term. As in, greenwashing, investors investing based on trends, or managers distributing funds inefficiently due to financially imperfect motives. For example, this would include a manager/investor investing into sustainability enhancing projects/more sustainable firms based on altruism without considering the value created/fundamental value of the firm. Borghesi et al. (2014) researched the various drivers that motivated company managers to make socially responsible investments. They found that female CEOs, younger CEOs, and managers who had donated to both Republican and Democratic parties were more likely to invest in Corporate Social Responsibility (CSR), and that media attention helped induce firms to increase CSR investments. These findings illustrated that not all all managers invest into the sustainability of their firm for the right reasons; if that were the case, then there would be no differentiation in CSR investment among the manager categories (sustainability trend among younger demographics), and politics and media would by no means play a

part in any probability to invest into CSR (moral hazard). The variation in manager motives to increase their firm's ESG scores creates informational asymmetries between investors and firms. Therefore, reasons for why firms obtain or pursue high ESG scores could be to the detriment of its financial performance. Instead of a firm's high sustainability indicating low risk to its cash flows, a bilateral relationship may arise where the cashflows also decrease in pursuit of sustainability, cancelling out the effects of a lower cost of capital on share prices. As a result, the relationship between ESG scores and returns likely suffers from endogeneity issues, of course not restricted to the cases above.

Perhaps the paper that paved the way for extensive research and focus surrounding ESG's, Orlitzy et al. (1998), titled "Corporate Social and Financial Performance: A Meta-analysis", explored the relationship between corporate social/environmental performance (CSP) and corporate financial performance (CFP). They analysed 52 previous quantitative inquiries into CSP's, yielding a sample size of 33,878 observations, and through a comprehensive analysis were able to confute the then mainstream claim that there was little generalizable knowledge about CSP. Specifically, they were able to establish a degree of certainty surrounding the CSP-CFP relationship which at that time was believed to be non-existent or insignificant among many business scholars, which today is fundamentally the equal to the relationship between ESGs and a firm's financial performance.

Some more recent papers have adopted more intricate methodologies to study the effects of ESG scores. For instance, Verheyden et al. (2016) conducted an in-depth analysis of the influence of ESG screening on return, risk, and diversification. Their study encompassed a comprehensive dataset of 85% of global public equities, and constructed Carhart (1997) four factor asset pricing portfolios by screening firms based on their ESG ratings. Specifically, they formed portfolios consisting of either the bottom x% or top x% of scores, allowing for an unambiguous comparative analysis between the portfolios' performances. The authors reported "an 'unequivocally positive' contribution to risk-adjusted returns when using a 10% best-in-class ESG screening" and that "the incorporation of ESG information contributes to better decision-making in every investment approach" (Verheyden et al., 2016, p.46). A possible problem with the analysis, however, may be the volume of observations given that 85% of global public equities are analyzed. It is essential to consider factors such as the sheer magnitude of observations, as it can bias parameter standard errors (SE) down, creating significant relationships which in fact might not be meaningful. Such considerations are of significance, as discussed in chapter 1 of this paper, given the varied and occasionally contradictory findings reported in research on this topic.

Luo (2022) examines the effects of ESG scores on stock returns of United Kingdom (UK) securities from 2003 to 2020. His approach uses a smaller sample compared to Verheyden et al. (2016), and thus is less prone to downward biased SEs of parameters due to sample size. The methodology Luo

(2022) entails multiple value weighted asset pricing models of portfolio returns, specifically, “the Fama–French (1993) three-factor model (FF3FM), the Carhart (1997) momentum-extended FF3FM, the Frazzini and Pedersen (2014) betting against beta-extended FF3FM, and the Asness et al. (2019) qualityminus-junk-extended FF3FM” (Luo, 2022, p.12). This allows a comprehensive controlled study of the practical implications of ESG investing for fund managers. The findings follow that firms with lower combined ESG ratings earn higher returns than firms with higher combined ESG ratings, and that ESG premiums are higher for environmental and social pillars compared to for combined ESG ratings. This contradiction to the previously discussed paper points to the fact that further research is needed in this field because theory as well as societal evolution point towards a positive and significant relationship.

CHAPTER 3 Data

3.1 Sample & Data Collection

Given this paper studies the relationship between ESG scores and firm returns, panel data (time-dimensional and firm cross-sectional) was selected from Eikon Refinitiv's database on all publicly traded equities. Specifically, data was collected on the firms' returns, combined ESG scores, respective E-S-G pillar scores, market capitalizations, price-to-book ratios, and Sharpe ratios. Since ESG scores are only calculated at the end of every fiscal year, observations were collected on a yearly basis for all the variables. The sample entailed an unbalanced panel (in terms of time) consisting of a general 11810 observations for each of the variables, worth 10 years of data (2013 – 2022), and 2117 different firms, spread across 14 different Western European. After cleaning the data however, the number of observations left in our sample is 10,695, a total of 1,972 firms. The selected countries were based on a political view of Western Europe, which on Eikon Refinitiv's database fell under Northern and Western European criteria (Austria, Germany, Belgium, France, Switzerland, Netherlands, Luxembourg, Republic of Ireland, United Kingdom, Finland, Norway, Iceland, Denmark, and Sweden). Specifically, the fixed effects model we utilize in our analysis omits time invariant variables due to collinearity. To see the variability in our relationship among regions we construct 17 separate universes. 14 different one specific to the respective countries, a Northern European universe, Western European Universe, and the general universe consisting of our entire data sample. Moreover, the data was extracted onto Microsoft Excel before being modified and imported into Stata for statistical analysis.

3.2 Variable Descriptions

3.2.1 Variables of Interest

Returns are the outcome variable of our research, measured as the change in share price of a firm within a single year. That is, the difference between the firm's opening share price on the first day of the year and the closing share price on the last day of the year, labelled *returns* in our regressions and defined as such in chapter 2.1.

The variable's purpose is to examine the impact of ESG scores on the financial performance of firms. Explicitly, equity performance due to the direct influence of the public on share prices and the global trend towards fostering sustainability. Other financial metrics such as revenues and return on investment (ROI) are not as directly prone to movement through the public's hand. Thus, the variable

aims to shed light on how ESG scores relate to a company's equity performance, taking into consideration the growing emphasis on sustainable practices worldwide.

The data itself is relatively normally distributed (as seen in Appendix B), and so no transformations were done to the variable. Interestingly, the original distribution of returns had a few extremely large positive outliers so returns of over 150% were dropped, this was also visually evident in scatter plots between returns and the various independent variables.

The *Combined ESG Scores and Pillar E-S-G* Scores are the main explanatory variables in our research. To be specific, the combined *ESG score* and the respective total individual pillar scores (*Environmental score*, *Social score*, and *Governance score*). The methodology used by Eikon Refinitiv for deriving the scores is discussed in chapter 2.2. Essentially, these scores represent metrics on a scale of 0-100 quantifying the level of sustainability demonstrated by a firm across each dimension, or in the case of the *Combined ESG Score,* holistically. Within the context of our study, the variables aim to measure whether the sustainability of a firm influences how it performs financially.

Since the variable is a rating metric between 0-100 no alterations of the data are strictly necessary. Despite this, an abnormality is observed with the firm environmental ratings. The rating of zero has twice the relative density of other ratings (Appendix B.1), and the irregularity is not mirrored by the remaining rating distributions. The magnitude of superfluous lower boundary ratings poses a statistical inconvenience, inconsistent parameter estimates due to intercept estimation becoming potentially inaccurate.

A rational assumption is that the reasoning behind this is attributed to the challenging and costly nature of implementing significant environmental sustainability measures for firms in certain industries, and possibly the inapplicability of environmental impact measures for certain industries. However, after an inspection of all the ESG score variables' summary statistics (in Table 1), it became apparent that the other ESG score variables failed to submit a single observation of zero (across over 44,000 observations). Given this finding and the sheer relative magnitude of zero values, the values were omitted from the sample as an error in data collection.

3.2.2 Control Variables

Market Capitalization is a continuous variable of the total value of a firm's equity (in euros), it measures the price of a firm's share multiplied by the number of outstanding shares (both values taken on the closing price of the last trading day of the respective observation's year). In essence, it is a

measure of the size of a public firm, controlling for factors like resource supremacy over smaller firms as the reason compelling the relationship between firm returns and ESG scores. After observing the distribution and summary statistics of the market capitalizations of firms in our sample it became apparent that distribution was exclusively positive and an exponentially decreasing function. To account for this, the natural logarithm of each observation was taken, bringing the distribution closer together while maintaining the integrity of the variable. Following that, the sample was visually cleaned for outliers through a scatter plot between the natural logarithm of market capitalizations and returns (the resulting distribution and relationship are presented in Appendix B).

In principle the *Sharpe Ratio* measures the risk adjusted return of an asset; the return earned by it over the risk-free rate divided by the standard deviation to adjust for risk (its algebraical form is available under Appendix A.3). Identically to above, it is a continuous variable, and it ensures that the relationship between ESG scores and returns is not influenced by the riskiness of a firm. As for the sample of data itself, the Sharpe ratios heavily centered around zero but with many firms still on either side of zero. Consequently, a signed logarithmic transformation of the variable was taken, followed by a scatter plot against returns to remove outliers. The final distribution and relationship can be seen in Appendix B.

Price-to-Book Ratio measures the public perception of the firm valuation. Whether it is over-valued or under-valuated in relation to its book value; a company's market capitalization divided by its book value of equity (as algebraically represented in Appendix A.4). It is also a continuous variable and aims to prevent the relationship between ESG scores and returns from being driven by differences in valuations and industry. In terms of distribution, the price-to-book ratio was analogous to the Sharpe ratio, hence the same transformation and following outlier exclusion followed. The final distribution and relationship can be seen in Appendix B.

3.3 Summary Statistics

Table 1: Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
returns	10695	0.074	0.360	-0.977	1.500
ESGC Score	10695	51.784	18.229	1.451	95.754
E Score	10692	49.581	25.983	0.020	99.094
S Score	10692	56.416	22.340	0.121	98.202
G Score	10695	54.065	22.368	0.462	98.733
lnmktcap	10695	21.670	1.730	16.013	26.757

signed ln pb ratio	10695	0.706	0.923	-2.482	3.892
sharpe ratio	10695	1.184	7.945	-24.821	34.826

Note. The following table contains the summary statistics of the all the variables from our data sample. For every respective variable, the number of observations, its mean, its standard deviation, its minimum observed value, and its maximum observed value are either calculated or gathered. It serves to give better insights into the variable data at hand, showing its abnormalities and defining characteristics.

As discussed in chapter 3.2 variable transformations were made and outliers were dealt with depending on their respective variable summary statistics, histograms, and scatter plots between the different independent variables and returns. Noteworthy is that there were quite a few extremely large returns registered by a few firms, one over 500%. Consequently, such observations were omitted from our sample, and the returns were cut at 150%. That seems plausible given that the worst performing return observation was -98% (no anomalies on the left tail) and the standard deviation is 0.36. Other than a few more extreme outliers among the control variables, there was nothing unique in the data.

CHAPTER 4 Methodology

To study the relationship between firm ESG ratings and returns I will construct a baseline regression model capable of analyzing the collected panel data (by time and firm-id). As explained in chapter 1, I enlist a similar methodology to Gavrilakis and Floros (2023), regressing identical variables. Specifically, the model type will be an Ordinary Least Square (OLS) regressions embedded with fixed effects and time-fixed effects. The inclusion of panel techniques differentiates our model from Gavrilakis and Floros (2023). From the scatter plots in Appendix C, it is visible that all our relationships examined are linear. Hence, modelling linear estimators seems appropriate. OLS is chosen because, satisfying the Classical Linear Regression Model (CLRM) assumptions, it produces the Best Linear Unbiased Estimators (BLUE), and its linear nature makes its parameters intuitive to both analyze and interpret.

Both fixed effects and time-fixed effects are employed by our baseline model. The employment of these panel techniques addresses endogeneity concerns, though it is important to highlight that a model may never completely eradicate this issue due to its inherent limitations. Specifically, fixed effects account for time-invariant variables not explicitly captured by our model; whist time-fixed effects account for time-variant variables not explicitly capture by our model. This is vital because not all variables can be directly controlled for by our model; without the panel techniques omitted variable bias would unequivocally make our estimates biased and inconsistent (endogeneity issue). In our case, fixed effects help control for the firm itself, by default encapsulating inherent factors like its industry and country, while time-fixed effects aid in controlling for factors that affect the average value of returns over time, such as policy changes and social trends. Given ESG investing popularity has grown partially due to policy reforms and sustainability movements, controlling for time-variant effects isolates the relationship between ESG ratings and returns from factors like social movements. In this way our model fundamentally estimates the ESG rating variables' parameters.

To satisfy the remaining four CLRM assumptions, we take the following measures. Firstly, we examine the mean of the residuals for each firm observation to be approximately zero. Next, to account for any potential heteroskedasticity, robust standard errors are employed akin to Arellano (1987). To tackle multicollinearity among variables, a correlation matrix table, and a subsequent variance inflation factor (vif) test are assessed between the variables (available in Appendix D, under Table 3 and 4). Furthermore, correlation between the residuals is accounted for by clustering standard errors according to firm id; appropriate given that we have firm panel data, so any correlation in residuals likely occurs within individual firms due to inherent characteristics shared within each firm.

These procedures satisfy the first four CLRM assumptions, making our model estimation approximately BLUE. To satisfy the final CLRM assumption, the residual distribution is visually inspected to be normally distributed through a histogram (visible in Appendix B, under Figure 3). The resulting regression equation follows:

$$returns_{firm,date} = \alpha + \beta_1\, ESG\_Score_{firm,date} + \beta_2\, \boldsymbol{Controls}_{firm,date} + \lambda_{date} + \mu_{firm} + \nu_{firm,date}$$

Where α is the constant outputted by the equation, $\beta ESG\_Score_{firm,date}$ is the effect of ESG scores on returns, $\beta Controls_{firm,date}$ the effect of the control variables on returns, λ_{date} the time-fixed effect (which in the actual regressions is imputed as a categorical control variable "$i.date$" as it is easier to specify the time invariant fixed effects), μ_{firm} the firm fixed effects, and $\nu_{firm,date}$ the remainder disturbance term. All the regressions are run on STATA®.

CHAPTER 5 Results & Discussion

5.1 Baseline Regressions

Our models used for the baseline regressions as well as the regressions to check for robustness were all OLS regressions. Therefore, a unitary increase in one of the non-transformed variables (ESG rating variables, sharpe ratio, and time-fixed effects) can be interpreted as causing a change in returns of the coefficient value of x. Moreover, for the transformed variables (firm market capitalization and price-to-book ratio), of which have had natural logarithms applied, a 1% increase in x causes a percentage change in returns of the coefficient value of x.

Table 2: Regressions results of our fully specified models, using our complete universe, measuring the relationship between the different ESG rating variables (Score: 0-100) and firm returns (%)

returns	(1)	(2)	(3)	(4)
ESGC_Score	-0.001***			
	(0.000)			
E_Score		-0.001***		
		(0.000)		
S_Score			-0.001**	
			(0.000)	
G_Score				0.000**
				(0.000)
lnmktcap	0.097***	0.098***	0.097***	0.096***
	(0.011)	(0.011)	(0.011)	(0.011)
signed_ln_pb_ratio	0.125***	0.124***	0.125***	0.125***
	(0.013)	(0.013)	(0.013)	(0.013)
sharpe_ratio	0.043***	0.043***	0.043***	0.043***
	(0.001)	(0.001)	(0.001)	(0.001)
Constant	-1.786***	-1.796***	-1.793***	-1.775***
	(0.241)	(0.241)	(0.242)	(0.241)
Number of obs	10695	10692	10692	10695
R^2 (within)	0.648	0.649	0.648	0.648
R^2 (between)	0.307	0.306	0.304	0.307
R^2 (overall)	0.472	0.475	0.472	0.469

Prob > F	0.000	0.000	0.000	0.000

*Note. This regression output table shows the parameter estimates of our full model specification using our entire universe. Specifically focusing on the relationship between firm returns and (1) firm combined ESG, (2) firm Environmental scores, (3) firm Social scores, and (4) firm Governance scores. The fully specified model includes all our controls as well as fixed effects and time-fixed effects (which have been omitted from our table but demonstrate appropriate values). The standard errors are noted in the brackets. *** p<0.010, ** p<0.050, * p<0.100*

Our model performed relatively well across all four ESG rating variables with on average the within R-squared being 65%, the between R-squared being 30%, and the overall R-squared being roughly 47%. The between R-squared is relatively low and is responsible for weighing down overall R-squared, however these results were expected given our inclusion of fixed effects was on the premise that returns vary among firms significantly due to the inherent unchanging characteristics that comprise each individual firm. As such, the within estimator is of most significance to us, indicating that our model elucidates the variability in returns.

Our variables of interest across the four regressions are the respective ESG rating variables of firms, and all are highly significant (even at the 1% level). The combined ESG score coefficient has a magnitude of 0.001 and a negative sign, meaning that for a unit increase in the combined ESG rating of a firm, and associated drop in its returns by 0.1% follows. The exact same details characterize the relationship between the environmental and social ratings of a firm and its returns, whilst the governance rating has no impact whatsoever (a significant coefficient of 0). In essence, excluding governance ratings, firms who observe an increase in ESG ratings are associated with a drop in returns during that same time-period. Practically, it seems that increasing the general sustainability of a firm negatively affects some part of its books or financial perception leading to a drop in returns. The same however cannot be said specifically for governance ratings for which investors do not seem to care about and do not subtract from a firm's financial health.

In conclusion, based on these preliminary results, our initial hypothesis- which expected to observe a positive relationship between the combined ESG score and firm returns- is rejected as we observe a negative effect. Despite this, the findings are consistent with previous literature which was previously highlighted as highly contradictory and sometimes of no significant use to investors. The latter is most definitely the case as governance ratings of firms are not associated with any changes in returns for a firm.

Secondarily, we hypothesized that the environmental rating of a firm would exhibit a stronger positive effect on relation to returns compared to the combined ESG score, social score, and governance score. However, our empirical findings did not support this claim. Surprisingly, we observed a negative effect of the environmental score on returns, with its effect being mirrored in sign, significance, and magnitude by both the combined ESG score and social score. Unexpectedly, the governance score demonstrated the most relatively favorable behavior wielding no effect over returns with a coefficient of zero. As a result, the second hypothesis is unequivocally rejected in favor of the governance score.

As for the three control variables, all demonstrate the expected sign and magnitude. The market capitalization of a firm is significantly positively associated with returns, so is the price-to-book ratio, and so is the sharpe ratio.

5.2 Robustness

We firstly check the robustness of our regression results to smaller subsamples of our universe. If the conclusions drawn from our baseline regressions are true and representative of Western Europe in a holistic sense, there should be no significant deviation in the results and conclusions drawn from regressions on smaller subsamples. Our fully specified models are regressed on country respective universes together with a Western European countries' universe and Northern Europe countries' universe. The regression results of these two models are visible in Appendix E under Table 6 and are compared to the baseline regressions from Table 2.

Analyzing Table 6 we can see that the findings from our baseline regressions are in doubt. For the most part, across all the different countries (other than Switzerland, Netherlands, UK, and Sweden) the ESG rating variables are completely insignificant, and so we must reject that ESG ratings effect firm returns in these countries. In the remaining countries, which produce some significant ESG rating variable coefficients, their characteristics are varied and the only interesting inferences they produce come from the regressions with the UK's universe and the Netherlands' universe. Within the UK the governance rating is significant and negatively associated with returns (magnitude of 0.001), a deviation to the conclusion made from Table 2 where the governance rating had no effect. Whilst, within the Netherlands, the combined ESG rating is significant, and in contrast to the conclusion made in Table 2, positive with a magnitude of 0.002. The latter observation leads to the observation that, whilst insignificant, there are a many positive coefficients- for example Denmark's coefficients are all positive. On the other hand, despite the largely insignificance coefficients, there is a visible similarity in terms of the magnitude of all the coefficients- regardless of country, they are generally extremely small. Overall,

however, our baseline conclusions cannot be representative of Western Europe given the deviation in results across countries. We revert to the fact that previous research is contradictory.

Additionally, we hypothesized that the Northern European countries of Western Europe would exhibit more positive ESG rating variable coefficients due to the perceptions of Northern Europe being the most sustainable region in Europe. From Table 6, we see that the coefficients for Northern Europe are all insignificant whilst the coefficients for Western Europe are all negative. As such, we cannot confirm that our hypothesis is reflected by our data. But we can also not fully reject it as inherently Western European firms experience a more negative effect over Northern European firms. We can say however, that our regressions are not robust to regional subsamples.

Furthermore, given the inconclusive nature of previous literature and the minimal magnitude of our coefficients thus far, we check for the robustness of our results to reduced forms of our baseline model from Table 2. Observing how the ESG rating variables' coefficients' change as our model becomes less informative, allowing for varying degrees of omitted variable bias (OVB), is key to understanding to what extent indirect factors essentially drive our relationship of interest. Accordingly, two reduced forms of our baseline models are run. Firstly, to observe the effects of trends and/or other time varying factors driving our relationships, baseline regressions lacking time-fixed effects are run. Secondly, in order to discern the remaining aspects of our relationship that exist as a result of indirect variables, the control variables and fixed effects are eliminated, leaving a univariate regression. The regression results of these two models in essence measure the robustness of our relationship rather than model. They are visible in Appendix E under Table 7, and a visual comparison of their coefficients to the coefficients from Table 2 are available in Appendix C, Figure 6.

Comparing Table 2 to Table 7 the change in the within, between, and overall average R-squared values are apparent, our fully specified regressions model the variance in returns with twice the accuracy of the time-fixed effects reduced model whilst the univariate regressions model nearly none of the variance in returns. Meaning, ESG ratings alone do not seem to coincide well with the variance of returns. The first thing to note about the coefficients in Table 7 is that all of the coefficients, bar the environmental rating variable in the univariate regression, are significant. Meaning that regardless of the degree of omitted variable bias ESG ratings are significantly associated with returns. Also, interestingly when removing time-fixed effects, the coefficients generally become more negative. Meaning, the relationship between ESG ratings and returns is positive over time, and not accounting for this bias's coefficients negatively. Theoretically, this is valid presuming a positive trend in sustainability with companies implementing ESG policies more ever-more-frequently. Finally, when removing all controls and panel techniques, we find that the combined ESG rating's, environmental ratings, and social rating's coefficients increase positively whilst the governance rating's coefficient increases negatively.

The positive increase indicates that, for those respective ESG ratings, the total effect introducing OVB into our model is positive, whilst for the governance rating it is negative.

In conclusion these reduced forms of our baseline regression highlight that the relationship between ESG rating variables and returns is not robust to OVB, and that in fact it is plausible the entire concept of ESG ratings being positively associated with returns is completely driven by endogenous variables.

5.2 Discussion

Principally, the results from our baseline model showed that ESG ratings other than the governance score exhibited a negative association to firm returns, with governance ratings exhibiting a null relationship. These findings are both similar and different to many previous studies highlighted by our paper.

Gavrilakis and Floros (2023), who shared identical control variables with our model, failed to find a significant association between ESG ratings and returns. This contrasts with our results as we did find a relationship. At that, a negative and significant one, indicating larger ESG ratings adversely affect stock returns. On the flip side however, our coefficients were small, which could also be interpreted as practically insignificant. A definite similarity with Gavrilakis and Floros (2023) was shared, however. The one significant relationship between ESG ratings and returns they obtained was negative, complementing the negative signs of all our ESG rating variable coefficients and reinforcing our results that ESG's negatively impact returns.

From a practical standpoint, the asset pricing model portfolios of Verheyden et al. (2016), illustrated that there was an unequivocal positive difference in the returns between the top 10% of combined ESG rated firms and bottom 10% of combined ESG rated firms. This contradicts our results as we observed a negative relationship between ESG rating variables and returns. Multiple theories explaining this difference are plausible. It may be that the US which was studied by their sample values ESG ratings more than Europe. But what seems more likely, given that we found the effects of the combined ESG score variable to positively increase when subject to OMV, is that their asset pricing model does not control for the appropriate endogenous variables.

Analogously to Verheyden et al. (2016), Luo (2022) ran identical portfolios utilizing more complex models and on a UK sample. They, in contrast to Verheyden et al. (2016), found identical results to us. A lower ESG rated firms made higher returns whilst higher ESG rated firms made lower

returns. The fact that the more complex model results in similar findings to us may indicate that the acknowledgement of endogenous variables shifts ESG rating's effects negatively. On the other hand, their study used a UK sample, and from our results we also established that on a country specific universe basis UK firms observed a negative relationship between ESG ratings and returns. Hence, it is unclear what is the main cause of differentiation in results among research.

Evident by now, our findings yield mixed results when compared to previous studies. Namely, they are both similar and different to previous research depending on the sample and model used. This however, rather ambiguously, can be interpreted favorably. A consequence of previous literature on our relationship of interest, between ESG ratings and firm financial performance, also exhibiting a contradictory and indecisive nature. Whelan et al. (2021) best illustrates this through their meta study on ESG ratings and financial performance, concluding that only 58% of corporate studies that examined the relationship between ESG ratings and financial performance (ROE, ROA, and stock performance) determined a positive relationship. Hence, collectively, our results are analogous to the broader literature available.

CHAPTER 6 Conclusion

6.1 Findings

In this thesis I have looked at the relationship between firms' ESG ratings and firms' stock returns in Western Europe. Previous literature has shown that we are not sure wither a relationship exists, but that it is slightly inclined towards truly being positive. Specifically for Western Europe, there are no famous studies, with most research focused on US markets; hence it is interesting to examine the relationship in Europe where the consensus is that ESGs are more dominant than in other markets. The contradictory findings coupled with a geographical lack of research underlies the importance of our research and why we explore whether firm ESG scores impact their stock performance in Western Europe.

To answer this question, across 14 different countries, a sample of 10,695 observation was collected, pertaining to a total of 1,972 firms, on their yearly stock returns, ESG scores, and control variables (market capitalizations, price-to-book ratios, and Sharpe ratios). To investigate and dissect the relationship multiple varying regressions were run on varying universes of our data. I ran baseline regressions, as well as reduced forms of my model, on my entire universe of data, and I also ran the baseline regressions on regionally segregated universes of that data. Our results revealed that the combined ESG scores, social scores, and environmental scores predicted negative returns for firms, whilst the governance scores had no effect. Furthermore, regionally, between Northern Europe and Western Europe, Northern Europe registered no effect, whilst Western Europe registered a negative effect. Noteworthy, the Netherlands registered one significant positive ESG rating coefficients. Finally, it was also revealed that when including omitted variables from our model, the combined ESG rating's, environmental ratings, and social rating's coefficients (bar the governance rating's coefficient which became more negative) became more positive. Despite the small magnitude of all our ESG variable related coefficients, our results are significant and the latter finding sheds light onto the possibility that, the reason many studies observe highly rated ESG portfolios predicting positive returns, may be due to its highly endogenous nature.

6.2 Implications

Given the mixed literature that exists an ultimate consensus is difficult to determine and frankly inappropriate. Results and conclusion vary depending on markets analyzed and methodologies enlisted, subjecting this paper to a no different fate.

It is valid to conclude though, that based on the constituents of this research, the relationship between ESG ratings and firm returns has been marginally negative. That is, over the last decade, Western European firms recording increases in their ESG ratings predicted a decrease in their returns. Unitarily, an increase in the combined ESG ratings, environmental ratings, and social ratings typically caused a 0.1% drop in returns. Depending on the reader's perspective, these effects may appear insignificant and possibly align more closely with the observed impact of governance ratings, which showed no predictive power for changes in returns. Consequently, our study contributes to the existing body of literature portraying ESG ratings as having either negative or inconsequential effects on firm returns. The presence of such a balanced and conflicting body of research may suggest that the true relationship between ESG ratings and returns may lie somewhere in-between (a valueless effect). Possibly driven by a regression to the mean and the inherent nature of model building that can be tailored to reflect an authors' viewpoints.

Regionally, all the significant coefficients were negative, with one positive exception for the Netherlands. This supports the general conclusion drawn by this paper that the effect of ESG ratings on returns is negative. A nuanced observation for researchers to keep in mind, however, is that despite being insignificant, many positive or neutral coefficients were observed across multiple countries. The consistency of these occurrences signals the possibility of a truly positive relationship, at the very least for certain countries. It may for example be a lack of data limiting their significance. Therefore, researchers should attempt to explore individual countries on a more detailed level and be aware for potential regional differences when building models.

With the latter in mind, we also concluded that the relationship between ESG ratings and returns is highly endogenous. When reducing the baseline regressions to univariate regressions we notice that the combined ESG score, environmental score, and social score coefficients increased. As such it is important to keep in mind that previous research may have inaccurately concluded that highly rated ESG portfolios perform well due to omitted variables biasing coefficients in a manner which portrays ESG ratings as positively influencing firm and portfolio returns. This bears particular significance and concern for investors, as they may be led astray by literature asserting that ESG portfolios outperform the market when, in fact, there is either a negative or non-existent relationship between ESG ratings and returns.

Thus, based on our research results, it would be inappropriate to assert that a variation in a firm's ESG ratings can directly cause a change in returns. It seems sounder to conclude that ESG ratings have no effect on firm returns, rather movements are endogenously explained. As such, if a firm's equity performance is of value, basing financial decisions on ESG ratings appears inaccurate and naive.

Accordingly, firms and investors should not consider ESG ratings of a firm when making policy changes or financial decisions, it should be an afterthought to filter between equally lucrative decisions and investment opportunities. In conclusion, future research should exercise prudence when investigating the relationship between ESG variables and firm returns, as it is possible that the observed results may not be solely attributable to ESG ratings.

6.3 Limitations

Unfortunately, data was not available for certain variables for certain firms across all 10 years. Consequently, our panel is unbalanced, resulting in a larger concentration of the data in our sample for more recent years. This can be a source for inconsistent parameter estimates and an increase in type 1 errors- especially for smaller samples such as the universes used for regressions in Table 6. This is problematic because we were forced to conclude that regional differences generally had no significant impact even though many countries produced positive coefficients. Had more data been available some of these coefficients may have been significant and lead to different inferences. Future research may look to estimate censored models on smaller samples of regional data.

Our paper also faces the constraint of being subject to a limited extent of practical deductions which can be extracted from our model. Being a standard linear regression model, the maximum we can derive from it is whether an association exists between our variables of interest. For instance, it would be overly ambitious to expect a definitive trading strategy based solely on a positive relationship. The same can be said to work vice versa though. Whilst a portfolio analysis approach would have been intriguing, the contradictory nature of previous research, including our own, emphasizes the need to first understand the fundamental relationship thoroughly, considering all its complex intricacies, such as the potential for endogeneity, before venturing into more complex models. Moving forward, additional research should focus on identifying the specific variables that predominantly influence the connection between ESG ratings and returns. As this paper attempts, laying the empirical foundations of any topic is critical when attempting to conduct more practical models because the ultimate findings lead to more accurate and meaningful conclusions. In essence, understanding the nuances of the relationship is crucial before attempting to draw concrete deductions from sophisticated models.

REFERENCES (APA)

admin. (2021, October 31). *What are ESG Metrics?* ESG | The Report. https://www.esgthereport.com/what-are-esg-metrics/

Alliance, G. S. I. (2017). *Global sustainable investment review 2016* [Report]. Global Sustainable Investment Alliance. https://apo.org.au/node/228336

Arellano, M. (1987). Computing Robust Standard Errors for Within-Groups Estimators. *Oxford Bulletin of Economics and Statistics*, *49*(4), 431–434.

Aroul, R. R., Sabherwal, S., & Villupuram, S. V. (2022). ESG, operational efficiency and operational performance: Evidence from Real Estate Investment Trusts. *Managerial Finance*, *48*(8), 1206–1220. https://doi.org/10.1108/MF-12-2021-0593

Asness, C. S., Frazzini, A., & Pedersen, L. H. (2019). Quality minus junk. *Review of Accounting Studies*, *24*(1), 34–112. https://doi.org/10.1007/s11142-018-9470-2

Böni, P., & Zimmermann, H. (2021). Are stock prices driven by expected growth rather than discount rates? Evidence based on the COVID-19 crisis. *Risk Management*, *23*(1–2), 1–29. https://doi.org/10.1057/s41283-021-00070-x

Borghesi, R., Houston, J. F., & Naranjo, A. (2014). Corporate socially responsible investments: CEO altruism, reputation, and shareholder interests. *Journal of Corporate Finance*, *26*, 164–181. https://doi.org/10.1016/j.jcorpfin.2014.03.008

Breitz, C., & Partapuoli, P. J. (n.d.). *How is ESG Affecting Stock Returns?*

Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. *The Journal of Finance*, *52*(1), 57–82. https://doi.org/10.1111/j.1540-6261.1997.tb03808.x

CFA, B. B. (n.d.). *ESG Investing Statistics 2023*. Bankrate. Retrieved May 15, 2023, from https://www.bankrate.com/investing/esg-investing-statistics/

De Lucia, C., Pazienza, P., & Bartlett, M. (2020). Does Good ESG Lead to Better Financial Performances by Firms? Machine Learning and Logistic Regression Models of Public Enterprises in Europe. *Sustainability*, *12*(13), Article 13. https://doi.org/10.3390/su12135317

Definition of Environmental, Social and Governance (ESG)—Gartner Finance Glossary. (n.d.). Gartner. Retrieved May 19, 2023, from https://www.gartner.com/en/finance/glossary/environmental-social-and-governance-esg-

Derwall, Guenster, Bauer, Koedijk, J., Nadja, Rob, Kees. (2019). *The Eco-Efficiency Premium Puzzle: Financial Analysts Journal: Vol 61, No 2*. https://www.tandfonline.com/doi/abs/10.2469/faj.v61.n2.2716?casa_token=j3pD040Ij8YAAAAA:DMJL5vxUysDUX8CA10HRQYdn51KE8DQRvaVgauo8GtFmKvTX96brTr75nGUVv7hbso72y0otWgOh

Discount Rate. (n.d.). Corporate Finance Institute. Retrieved June 29, 2023, from https://corporatefinanceinstitute.com/resources/valuation/discount-rate/

ESG performance, herding behavior and stock market returns: Evidence from Europe | SpringerLink. (n.d.). Retrieved May 15, 2023, from https://link.springer.com/article/10.1007/s12351-023-00745-1

ESG scores and cost of debt—ScienceDirect. (n.d.). Retrieved May 19, 2023, from https://www.sciencedirect.com/science/article/pii/S0140988322003383?casa_token=I8dyfpQtojkAAAAA:7kAr7MswE248dwXjf3mXFXl42ClXenNyCfisdY706KlnrrQ8C9AnZpZEze0PKzrsYsa3Vkokwg

Eugene, James, F. F., D. MacBeth. (1973). *Risk, Return, and Equilibrium: Empirical Tests | Journal of Political Economy: Vol 81, No 3*. https://www.journals.uchicago.edu/doi/abs/10.1086/260061

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3–56. https://doi.org/10.1016/0304-405X(93)90023-5

Ford, J. M., Gehricke, S. A., & Zhang, J. E. (2022). Option traders are concerned about climate risks: ESG ratings and short-term sentiment. *Journal of Behavioral and Experimental Finance*, *35*, 100687. https://doi.org/10.1016/j.jbef.2022.100687

Frazzini, A., & Pedersen, L. H. (2014). Betting against beta. *Journal of Financial Economics*, *111*(1), 1–25. https://doi.org/10.1016/j.jfineco.2013.10.005

Girerd-Potin, I., Jimenez-Garcès, S., & Louvet, P. (2014). Which Dimensions of Social Responsibility Concern Financial Investors? *Journal of Business Ethics*, *121*(4), 559–576. https://doi.org/10.1007/s10551-013-1731-1

Gordon, M. J. (1959). Dividends, Earnings, and Stock Prices. *The Review of Economics and Statistics*, *41*(2), 99–105. https://doi.org/10.2307/1927792

Graham, B. (1939). Review of The Theory of Investment Value [Review of *Review of The Theory of Investment Value*, by J. B. Williams]. *Journal of Political Economy*, *47*(2), 276–278.

How Do I Calculate Total Return On Investment? (2023, January 14). Yahoo Finance. https://finance.yahoo.com/news/calculate-total-return-investment-140017092.html

Kliment, Dolák, P., Pavel. (2022, December 6). *Two thirds of CEE banks already have an ESG strategy for real estate financing—KPMG Czech Republic*. KPMG. https://kpmg.com/cz/en/home/media/press-releases/2022/11/two-thirds-of-cee-banks-already-have-an-esg-strategy.html

Luo, D. (2022). ESG, liquidity, and stock returns. *Journal of International Financial Markets, Institutions and Money*, *78*, 101526. https://doi.org/10.1016/j.intfin.2022.101526

Orlitzky, Schmidt, Rynes, M., Frank L. ,. Sara L. (2003). *Corporate Social and Financial Performance: A Meta Analysis*. https://journals.sagepub.com/doi/abs/10.1177/0170840603024003910

Performance Metrics Examples and Definition. (2023, May 10). Sage Advice US. https://www.sage.com/en-us/blog/glossary/what-are-performance-metrics/

Refinitiv® Datastream®. (n.d.).

Sustainable investing—Country Sustainability Ranking. (n.d.). Robeco.Com - The Investment Engineers. Retrieved May 15, 2023, from https://www.robeco.com/en-int/glossary/sustainable-investing/country-sustainability-ranking

Terminal Growth Rate. (n.d.). Corporate Finance Institute. Retrieved June 29, 2023, from https://corporatefinanceinstitute.com/resources/valuation/what-is-terminal-growth-rate/

The Evolution of ESG Investing—MSCI. (n.d.). Retrieved May 19, 2023, from https://www.msci.com/esg-101-what-is-esg/evolution-of-esg-investing

Verheyden, T., Eccles, R. G., & Feiner, A. (2016). ESG for All? The Impact of ESG Screening on Return, Risk, and Diversification. *Journal of Applied Corporate Finance*, *28*(2), 47–55. https://doi.org/10.1111/jacf.12174

What are stock market returns? (n.d.). MoneySense. Retrieved May 15, 2023, from https://www.moneysense.ca/glossary/what-are-stock-market-returns/

Whelan, T., Atz, U., & Clark, C. (n.d.). *ESG AND FINANCIAL PERFORMANCE:*

Wong, W. C., Batten, J. A., Ahmad, A. H., Mohamed-Arshad, S. B., Nordin, S., & Adzis, A. A. (2021). Does ESG certification add firm value? *Finance Research Letters*, *39*, 101593. https://doi.org/10.1016/j.frl.2020.101593

Wu, Y., & Borovkova, S. (n.d.). *ESG considerations in your factor strategies*.

Yao, T. (n.d.). *ESG Education and Job Choice*.

APPENDIX A: Equations

A.1 Gordon Growth Model (GGM)

$$P = \frac{D_0(1+g)}{r-g}$$

P represents the intrinsic value or fair value of the stock

D_0 is the most recent dividend paid out by the company

g is the expected constant growth rate of dividends

r is the required rate of return or discount rate

A.2 Discounted Cash Flow Model (DCF)

$$DCF = \frac{CF_1}{(1+r)^1} + \frac{CF_2}{(1+r)^2} + \cdots + \frac{CF_n}{(1+r)^n} + \frac{TV}{(1+r)^n}$$

$$TV = \frac{CF_{n+1}}{r-g}(1+g)$$

DCF represents the discounted cash flow, which is the estimated present value of future cash flows

CF represents the cash flow in each period, typically annual or quarterly

r is the discount rate, representing the required rate of return or the rate at which the future cash flows are discounted

n represents the number of periods (usually years) over which the cash flows are projected

TV denotes the terminal value, which represents the estimated value of the company beyond the projected period

g is the expected constant growth rate of the cash flows

A.3 Sharpe Ratio

$$s = \frac{r_a - r_f}{\sigma_a}$$

s represents the Sharpe ratio calculated

r_a represents the return of the asset over a period

r_f represents he risk free rate of return over that time

σ_a represents the standard deviation of the return of the asset

A.4 Price-to-Book Ratio Per Share

$$PB\ Ratio\ Per\ Share = \frac{Market\ Share\ Price}{Book\ Value\ of\ Equity\ Per\ Share}$$

APPENDIX B: Distributions

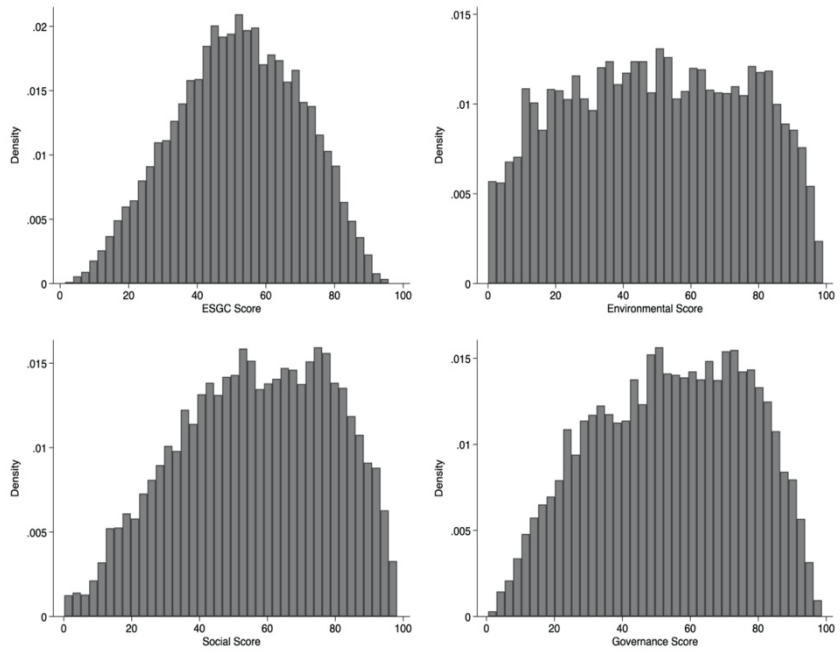

Figure 1. Histograms of the distributions of the ESG rating variables of the complete data universe

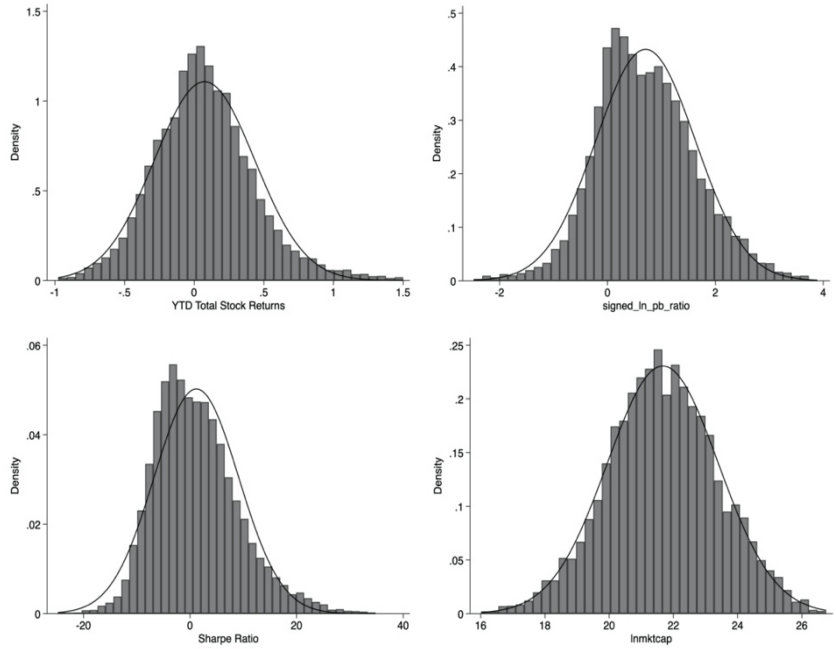

Figure 2. Histograms of the distributions of the dependent variable and control variables of the complete data universe

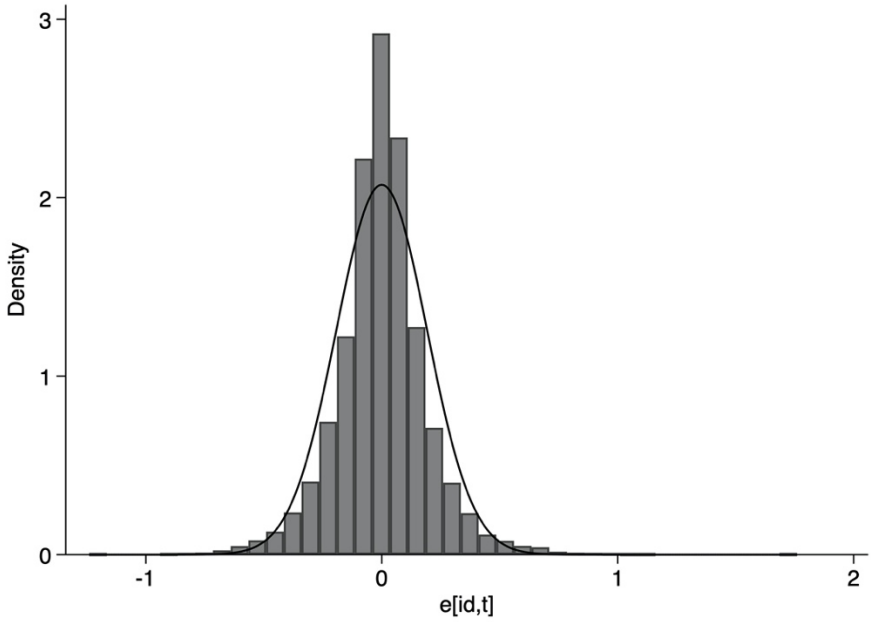

Figure 3. Histogram of the distribution of our baseline model's residuals when using the combined ESG rating as the sole ESG rating variable of our regression

APPENDIX C: Plots

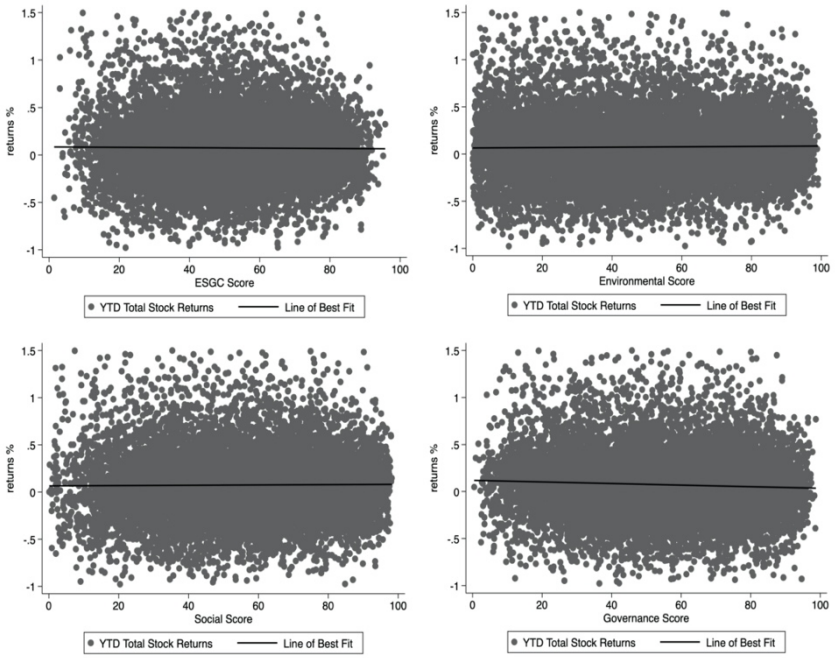

Figure 4. Scatter plots of returns on ESG rating variables for our complete universe

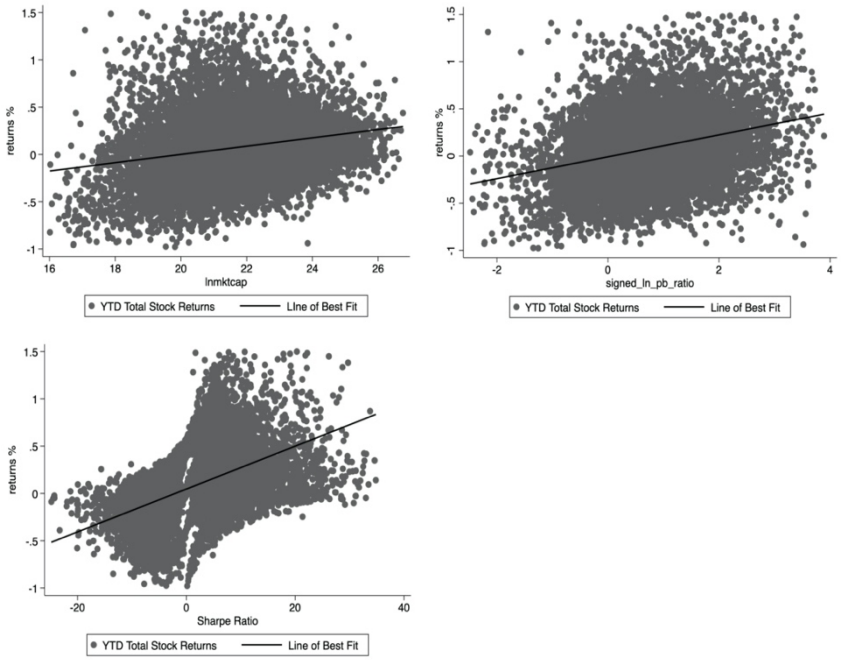

Figure 5. Scatter plots of returns on the Control variables for our complete universe

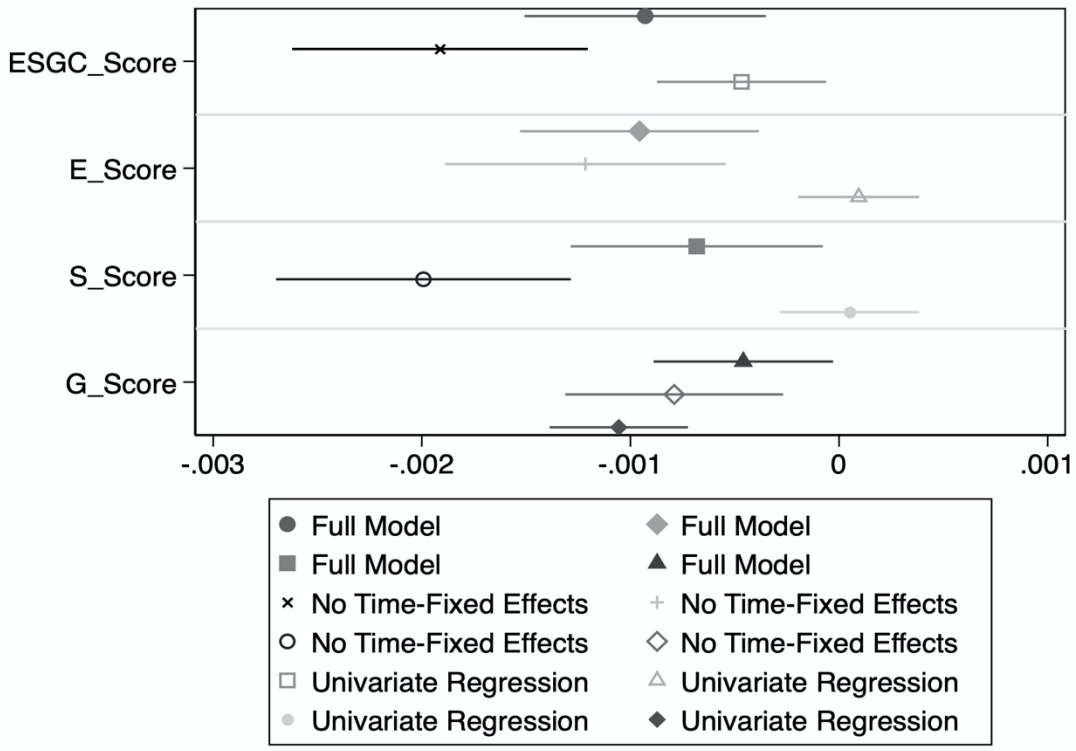

Figure 6. Coefficient plot comparing the ESG rating variables' coefficients for a fully specified model, a fully specified model less time-fixed effects, and a univariate regression using our complete universe

APPENDIX D: Data & Model Building Tests

Table 3: Matrix of Correlations

Variables	(1)	(2)	(3)	(4)	(5)
(1) returns	1.000				
(2) ESGC_Score	-0.010	1.000			
(3) lnmktcap	0.210	0.476	1.000		
(4) signed_ln_pb_ratio	0.298	0.026	0.205	1.000	
(5) sharpe_ratio	0.500	-0.004	0.184	0.189	1.000

Note. The following table contains the respective correlations between each of our variables (dependence variable and independent variables, including control variables as well). It shows whether a relationship between variables exists at a first glance and highlights potential multicollinearity issues.

Table 4: Variance inflation factor

	VIF	1/VIF
ESGC Score	6.695	0.149
S Score	3.87	0.258
E Score	3.111	0.321
G Score	2.143	0.467
lnmktcap	1.884	0.531
signed ln pb ratio	1.151	0.869
sharpe ratio	1.078	0.927
Mean VIF	2.847	

Note. This table illustrates the results of a variance inflation factor (vif) test run after a standard regression of our dependent variable on all our independent variables. The test is the standard practice test conducted to identify multicollinearity between independent variables. It is done due to correlation coefficients from table 2 indicating possibly multicollinear variables.

Table 5: Hausman (1978) specification test between fixed effects and random effects

	Coef.
Chi-square test value	698.471
P-value	0

Note. This table shows the Hausman test between two panel models, one specified with panel fixed effects and the other with panel random effects. This is the standard test for statistically classifying whether a panel of data should be specified with fixed effects or random effects. It statistically checks which model's coefficients are consistent and efficient, concluding which technique better represents the data at hand.

APPENDIX E: Robustness Tests

Table 6: Full model regression coefficient results for the different ESG rating variable (Score: 0 - 100) parameters estimating their relationship with returns (%), run on country specifics universes, a Northern European countries' universe, and a Western European countries' universe

	Country Universe				
returns	(Austria)	(Germany)	(Switzerland)	(France)	(Belgium)
ESGC_Score	0.002	-0.001	-0.003**	0.000	-0.002
	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)
E_Score	0.000	-0.001	-0.002*	0.000	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
S_Score	0.001	0.000	-0.002***	0.000	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
G_Score	0.001	-0.001	-0.001	0.000	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
returns	(Netherlands)	(UK)	(Rep. Ireland)	(Denmark)	(Sweden)
ESGC_Score	0.002*	-0.002***	0.002	0.000	-0.001
	(0.001)	(0.001)	(0.003)	(0.002)	(0.001)
E_Score	0.001	-0.002***	0.002	0.001	-0.002**
	(0.001)	(0.000)	(0.003)	(0.002)	(0.001)
S_Score	0.001	-0.001**	0.003	0.001	-0.001
	(0.001)	(0.001)	(0.003)	(0.002)	(0.001)
G_Score	0.001	-0.001*	0.000	0.000	0.000
	(0.001)	(0.000)	(0.002)	(0.001)	(0.001)
returns	(Norway)	(Finland)	(Iceland)	(N. Europe)	(W. Europe)
ESGC_Score	0.000	-0.001	-0.005	-0.001	-0.001***
	(0.002)	(0.001)	(0.010)	(0.001)	(0.000)
E_Score	0.001	-0.002	0.007	-0.001	-0.001***
	(0.002)	(0.001)	(0.009)	(0.001)	(0.000)
S_Score	0.000	0.000	0.005	0.000	-0.001**
	(0.002)	(0.001)	(0.022)	(0.001)	(0.000)
G_Score	0.000	0.001	0.003	0.000	-0.001***
	(0.001)	(0.001)	(0.003)	(0.001)	(0.000)

Note. This regression output table shows the parameter estimates for our ESG rating variables.

*Specifically, four fully specified models are run (one for each ESG rating variable) on 14 country exclusive universes, a Northern European countries' universe, and on a Western European countries' universe (the segregation of Northern and Western countries from within Western Europe). Notably, the country Luxembourg has been omitted from the table due to a lack of firm observations limiting the calculation of standard errors and t-statistic. Additionally, the control variables have been manually removed from the table (to not clutter the table) because they are all generally significant, meet the expected magnitudes and signs as discussed in Chapter 5.1, and our main analytical concern is with the comparison of ESG rating parameter estimates between countries. The standard errors are noted in the brackets. *** p<0.010, ** p<0.050, * p<0.100*

Table 7: Regressions results, using our entire universe, measuring the relationship between the different ESG rating variables (Score: 0-100) and firm returns (%) through a univariate model and a fully specified model less panel time-fixed effects

	Univariate Models				Full Models less time-fixed effects			
returns	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ESGC_Score	0.000**				-0.002***			
	(0.000)				(0)			
E_Score		0.000				-0.001***		
		(0.000)				(0.000)		
S_Score			0.000***				-0.002***	
			(0.000)				(0.000)	
G_Score				-0.001***				-0.001***
				(0.000)				(0.000)
lnmktcap					0.175***	0.173***	0.177***	0.172***
					(0.014)	(0.014)	(0.014)	(0.014)
signed_ln_pb_ratio					0.193***	0.195***	0.193***	0.195***
					(0.016)	(0.016)	(0.016)	(0.016)
sharpe_ratio					0.016***	0.016***	0.016***	0.016***
					(0.000)	(0.000)	(0.000)	(0.000)
Constant	0.095***	0.066***	0.068***	0.126***	-3.781***	-3.781***	-3.800***	-3.758***
	(0.011)	(0.008)	(0.01)	(0.01)	(0.294)	(0.296)	(0.295)	(0.295)
Number of obs	10695	10692	10692	10695	10695	10692	10692	10695
R^2 (within)	0.017	0.011	0.091	0.017	0.376	0.375	0.377	0.374
R^2 (between)	0.024	0.022	0.030	0.003	0.187	0.186	0.185	0.188
R^2 (overall)	0.000	0.000	0.000	0.003	0.184	0.184	0.186	0.182
Prob > F					0.000	0.000	0.000	0.000
Prob > chi2	0.024	0.525	0.764	0.000				

*Note. This regression output table shows the parameter estimates for two reduced forms of our main model (1-4 is a univariate model, and 5-8 is our fully specified model less the time-fixed effects) using our entire universe. For both models there are four regressions, representing that each regression determines the relationship to firm returns (%) for one of the ESG rating variables our paper studies. The standard errors are noted in the brackets. *** p<0.010, ** p<0.050, * p<0.100*