Norwegian School of Economics
Bergen, Fall 2021

ESG and stock market performance

An Empirical Study of the Link between ESG Performance and Stock Performance of Scandinavian Companies

Omar Båfjord Ismaili and Adrian Ekroll Kjøsnes

Supervisor: Konrad Raff

Master thesis, Economics and Business Administration
Majoring in Financial Economics and Business Analysis and Performance Management

NORWEGIAN SCHOOL OF ECONOMICS

Acknowledgements

This thesis marks the completion of our Master of Science at the Norwegian School of Economics, majoring in Financial Economics and Business Analysis and Performance Management.

We want to thank our supervisor, Konrad Raff, for providing guidance through the process and helpful discussions. His insights and encouragements have been valuable. Furthermore, we want to thank our friends and family for feedback and continuous support throughout the semester. Last but not least, we thank each other for a great partnership.

Norwegian School of Economics

Bergen, 2021, December 20^{th}

Omar Båfjord Ismaili

Adrian Ekroll Kjøsnes

Abstract

This thesis investigates the relationship between environmental, social, and corporate governance (ESG) performance and stock performance in Scandinavia between 2011-2020. We analyze the difference in stock performance between companies with high and low ESG performance. To measure ESG performance, we apply various ESG ratings from Refinitiv. Using a long-short zero investment strategy, we examine differences in stock performance between high and low ESG rated companies. By employing Fama-French three-, four- (Carhart) and five-factor models, we control for possible different risk exposures between the portfolios. The portfolios based on ESG score show a neutral relationship. The neutral relationship is consistent when controlling for ESG controversies, the company's media exposure related to ESG incidents. Amongst the three ESG dimensions, we also find a neutral relationship with portfolios based environmental and social scores. However, screening on governance score, the strategy leads to abnormal returns, indicating a positive relationship.

Keywords – CSR, ESG, Stock Performance, Environmental, Social, Governance

Contents

List of Tables

1 Introduction

The demand for sustainable investments has increased rapidly over the last decade. As of 2020, global sustainable investments reached more than 35$ trillion, compared to 22$ trillion in 2016. Thus, representing 35,9% of total assets under management in 2020 (GSIA, 2021). Sustainable investing is commonly referred to as ESG investing. The acronym ESG (environmental, social and governance) has emerged to specify business sustainability. ESG refers to how businesses incorporate environmental, social, and governance concerns into their operations, and it is gaining traction in research, investment, and media coverage. Nicolai Tangen, CEO of Norges Bank Investment Management, expressed his thoughts on the increasing importance of ESG implying companies that do not embrace ESG will eventually disappear (Taraldsen, 2021).

Several studies have examined the relationship between ESG and stock performance by examining ESG-screened portfolios. Traditional portfolio theory (Markowitz, 1959) suggests that any ESG screen diminishes returns since it limits investors' options. However, according to literature, ESG screens have historically given mixed results. Studies conducted by Halbritter and Dorfleitner (2015) and Statman and Glushkov (2009) found a neutral effect on stock performance and studies by Hong and Kacperczyk (2009) and Renneboog et al. (2008) found a negative effect of ESG screens on performance. While the studies by Edmans (2011), Derwall et al. (2005) Kempf and Osthoff (2007) and Eccles et al. (2012) found positive effects, indicating ESG performance may lead to a positive impact on stock performance.

This thesis examines the relationship between ESG performance and stock performance of companies listed on the Scandinavian[1] stock exchanges, using a sample of 113 companies. In order to study this relationship, we used ESG scores from Refinitiv[2] as a proxy for ESG performance. ESG scores are an aggregated evaluation of the firm's ESG performance on three pillars: environmental, social and governance. Refinitiv is one of the leading ESG data providers, with coverage of over 80% of market capitalization in the world (Refinitiv,

[1] Norway, Sweden, Finland and Danmark
[2] Formerly known as Thomson Reuters

2021). To ensure an objective evaluation of ESG performance, Refinitiv relies purely on publicly available information. Thus, ensuring that our analysis uses a proxy for ESG performance that accurately reflects the information available to ESG investors.

In the analysis, we divided the companies into quintiles based on ESG performance. Furthermore, we construct portfolios using the long-short strategy. The long-short strategy[3] is a net-zero investment strategy, buying the top quintile portfolio and selling short the bottom quintile portfolio. The overall ESG scores, ESG and ESGC, and the individual environmental, social, and governance pillars are considered as measures of ESG performance. Further, we rebalance our portfolios yearly, with a one-month lag after new ESG-scores are released. By lagging we ensure that the information is available to the market and prevent look-ahead bias. To measure performance and control for differences in risk exposure between the portfolios, we employ the Fama-French three-factor model, the Carhart four-factor model, the Fama-French five-factor model with and without momentum (Fama and French, 1993; Carhart, 1997; Fama and French, 2015).

Our results show no statistical difference in stock performance between companies with high- and low ESG performance, when using overall ESG scores and the individual environmental and social scores. Thus, indicating a neutral relationship between ESG performance and stock performance. However, based on the governance scores, the long-short portfolio generates abnormal returns. We also find a difference in exposure to systematic risk between companies with high and low ESG performance, when sorting the portfolios on ESG score and social score. ESGC based portfolios shows less robust differences. The difference in beta implies that companies with high ESG performance are less exposed to systematic risk.

The thesis contributes to the existing empirical literature on the relationship between ESG performance and stock performance. By utilizing data from Scandinavian markets, our research differs from the majority of literature, which focuses primarily on US markets. Further, by examining the relationship using environmental, social, and governance scores, we examine how performance in each ESG dimension affects stock

[3]The long-short strategy is also referred to as the "zero-investment strategy" or the "high-low strategy" (Alexander, 2000).

performance.

The remainder of the thesis is structured as follows: The second section examines the theoretical and empirical links between ESG and stock performance, as well as ESG rating disagreement. Our hypothesis is presented in the third section. Section 4 describes our sampling procedures, the final data sample, and the methodology for calculating ESG ratings. Section 5 describes the empirical approach used in the analysis, and Sections 6 and 7 present our findings and additional discussion. Lastly, Section 8 brings the thesis to a close with the conclusion.

2 Theory and literature review

The following section will present the theoretical and empirical motivation behind the thesis and seek to explain the relationship between ESG and stock performance. Firstly, it will present the relevant economic theories. Secondly, we present relevant research regarding the relationship between ESG and stock performance.

2.1 Shareholder theory

> "There is one and only one social responsibility of business – to ... increase its profits" – Milton Friedman

According to Milton Friedman (1970), the sole purpose of business is to maximize shareholder value. The Friedman doctrine serves as the foundation for the shareholder theory. Shareholders are considered as the only group to whom the company is socially responsible. As a result, the goal of the company is to maximize shareholder profits. However, Milton Friedman's viewpoint is frequently perceived as being less nuanced than it is. While the title of Friedman's doctrine from 1970, "The Social Responsibility of Business is to Increase Its Profits", suggests that his stance is clear. Friedman acknowledges individuals having social responsibilities beyond the scope of profits. By maximising profits, businesses fulfil their social responsibility giving individuals maximum flexibility to fulfil their individual social responsibilities (Friedman, 1970). Participation in corporate social initiatives should be done solely by shareholders, rather than indirectly through managers acting on their behalf. Any engagement in philanthropy or activities that are not associated with earning profits will result in diminishing returns. Thus, as Friedman (1970) ultimately suggests, considering additional stakeholders is value-destroying and shareholders should be the only priority.

2.2 Stakeholder theory

The Business Roundtable[4] announced in 2019 a change to the "Purpose of a Corporation" from the traditional shareholder perspective to the stakeholder perspective (Business

[4]The Business Roundtable is a nonprofit lobbyist association, consisting of CEO's from major American companies.

Roundtable, 2019). The stakeholder theory was first presented by R. Edward Freeman in 1984 as a response to Friedman's stakeholder theory. To be successful, Freeman believes that businesses must create value for all stakeholders, including shareholders. By including the needs and interests of all stakeholders in their business models, businesses may experience improved financial performance. Similarly, businesses may experience reduced financial performance by failing to maintain stakeholder relations. Shareholder value is only maximized by considering all stakeholders over time (Freeman and Phillips, 2002). Ultimately, the only way to ensure long-term financial success.

2.2.1 Pieconomics

Pieconomcs offers a new perspective on the trade-off between shareholder value and social responsibility. In Edmans' theoretical framework, the total shareholder and stakeholder value are represented as a pie (Edmans, 2021). Unlike the traditional stakeholder and shareholder theories, the size is not fixed. Actions impact both shareholder and stakeholder value. However, one does not necessarily come at the expense of the other. Both can be improved by initiating processes and actions that increase the size of the pie, which Edmans refers to as the pie-growing mentality (Edmans, 2021, p. 26). Unlike the traditional view of a "pie splitting" mentality, where shareholder and stakeholder value is created at the expense of the other. Moreover, Edmans (2021) states that profits are a by-product of contributing to society. Social value, rather than profits, should be the main driver of businesses and also the main focus. The main difference between Pieconomics and the traditional theories is recognizing profits as an outcome of social value, not a primary goal.

2.3 Review of empirical literature

2.3.1 CSR, ESG and SRI

Sustainability is often referred to as "Corporate Social Responsibility" (CSR) in research, and both "Socially Responsible Investing" (SRI) and "Environmental, Social, and Governance" (ESG) in its implementation in finance. While the expressions partly differ, all expressions are based on the same basic principles and revolve around the three ESG factors. When presenting the related literature, the commonly used expression in the

paper is referred to, but throughout this thesis ESG is used. ESG performance is often measured by ESG scores from different rating providers, e.g., Bloomberg, MSCI, Refinitiv and KLD Analytics. The ESG score is a numerical aggregation of perceived performance across various environmental, social, and governance topics. Other individual performance measures are also commonly used in literature to investigate the relationship between specific topics and financial performance. The use of specific measures is dominant in governance literature, where extensive research has been conducted prior to the recognition of social and environmental factors in research.

2.3.2 ESG and Stock Performance

Several studies examine the relationship between the companies financial and social performance based on the ESG score. Kempf and Osthoff (2007) found that implementing a long-short investment strategy of buying high SRI rated companies and selling low rated SRI companies, generates high abnormal returns for the period 1992-2004. Kempf and Osthoff (2007) measured SRI performance using ESG ratings from KLD Research & Analytics. The same relationship is found by Statman and Glushkov (2009), examining the relationship between 1992 and 2007 using KLD scores and found that a high-low strategy gives significant positive abnormal returns. Consistent with this, Eccles et al. (2014) found similar out-performance by high ESG score companies using a high-low strategy. Eccles et al. (2014) combined ESG ratings from ASSET4 and Sustainable Asset Management (SAM) together with personal research and interviews to divide their sample into high or low ESG companies in their research.

Oppositely, Borgers et al. (2013) show that the outperformance of high ESG score companies disappears after the initial sampling period from Kempf and Osthoff (2007). Arguing the positive abnormal returns were found due to market underreactions. Abnormal returns are created when the markets corrects the valuation of initially undervalued intangible ESG benefits as they become tangible through increased earnings Bénabou and Tirole (2010). Borgers et al. (2013) argue if such mispricing is happening, the earnings announcements of high ESG firms will surprise positively. Borgers et al. (2013) found such a relationship in their sample until 2004. However, the relationship disappears between 2004-2009, suggesting that the market succeeds in valuing the ESG

performance, and the mispricing disappears. Similarly, Halbritter and Dorfleitner (2015) finds a neutral relationship over the period 1991 to 2012 using ESG ratings from ASSET4, Bloomberg and KLD.

Krüger (2015) studied short-term stock reactions to ESG events. The event study discovered that the stock market reacted negatively to businesses' ESG initiatives. Krüger (2015) found that the market responded negatively to negative events and weakly negatively to positive events. The effect of the negative events is unsurprising, as they often imply negative cash flows. Further, he argues the negative effect on positive initiatives is due to agency concerns and that the leadership's actions therefore are to the detriment of shareholder value. However, Krüger (2015) find that stock prices rise when the positive news are related to managerial efforts offsetting prior social irresponsibility.

Other studies have used more specific screens to examine whether stakeholder relations impact stock performance. According to Edmans (2011), companies with stronger employee satisfaction had higher risk-adjusted returns between 1984-2009. Subsequently, they also exhibit higher earnings announcement returns and higher long-term earnings surprises. Edmans (2011) presents extensive evidence that the stock market fails to price the intangible assets of companies with strong employee relations. Statman and Glushkov (2009) and Kempf and Osthoff (2007) found that a high-low strategy based on employee relations ratings generated significant abnormal returns, thus being consistent with the results of Edmans (2011). Additionally, Servaes and Tamayo (2013) find a positive relationship between firm value and customer awareness. However, they also found that companies with high customer awareness is penalized more when ESG concerns are present.

The evidence from environmental screening investment strategies are mixed. Derwall et al. (2005) measure environmental performance using Innovest ratings as a proxy for corporate eco-efficiency. Their results show that companies with high eco-efficiency outperform those with low eco-efficiency over an eight-year period. However, the results lack robustness, only being significant at the 5% level. The authors also back-filled the Innovest rating data, extending the time-period by two years making the research suspect to look-ahead-bias. Further, Halbritter and Dorfleitner (2015) find contradictory results

when using KLD environmental scores. They find an over-performance for high ranked companies using a high-low strategy for the time-period 1990-2001. However, the strategy obtains insignificant alpha for 2002-2012. The neutral difference is consistent with Kempf and Osthoff (2007) that finds a neutral difference in stock performance with a high-low strategy based solely on the KLD environmental score.

Studies[5] using governance screens find similar evidence as the environmental screens. Gompers et al. (2003) finds that firms with strong stakeholder rights exhibit yearly risk-adjusted stock returns that are 8,5% higher than those with weak shareholder rights in the period 1990 to 1999. Their findings also show that companies with weak governance practices persistently underperform the market. However, a study by Core et al. (2006) tests for causality over the same period by seeing whether the stock market is surprised by the poor operating performance of weak governance firms. The results indicate neither analysts' forecast errors nor earnings announcement returns show no sign that the underperformance surprised the market. Furthermore, Bebchuk et al. (2013) extends the sample period of Gompers et al. (2003) to cover the period between 1990 and 2008. Their results show that the abnormal returns between 2000 and 2008 were insignificant. Indicating the effect has vanished since the initial sample period of Gompers et al. (2003).

Hong et al. (2012) argue that CSR activities rise with firms' financial performance, but not oppositely. They discover that goodness spending is more sensitive to financial slack compared to capital and R&D expenditure. Hong et al. (2012) argue firms "do well" and build up financial slack, which allows them to "do good" by engaging in CSR activities. As a result, the less-constrained firms spend more on goodness. The discovery also highlights the literature's concerns about reverse causality, whether firms "do well by doing good" or the oppositely.

The relationship between ESG and firm risk is also covered in the literature. Bénabou and Tirole (2010) observed that firms with stronger ESG profiles may have different systematic risk exposures due to their resilience during crisis periods or because

[5]Corporate governance is prominent in finance research. Research involves topics as board composition, directors compensation, governance practises, shareholder rights and diversity. The mentioned studies have been chosen due to their similarities in terms of methodology and their use of an aggregated measure of corporate governance performance.

of a specific ESG risk factor. Furthermore, Lins et al. (2017) find that high ESG firms performed better during the 2008–2009 financial crisis, supporting the resiliency argument. Albuquerque et al. (2019) present a theory in which firms with high ESG performance face relatively less price elastic demand as a result of ESG being a product of differentiation strategy. Their finding show that companies with high CSR rating have a lower cost of capital due to lower systematic risk.

2.3.3 ESG rating weaknesses

ESG rating agencies provide a third-party assessment of firms' perceived ESG performance. Rating agencies collect and assess information on various environmental, social and governance issues, resulting in an ESG numeric score or ESG rating for the specific firm. Today, the ESG rating industry consists of a few large providers, which have a large influence on both sustainable investments and the literature concerning sustainable investments (Berg et al., 2019).

Research show significant variations in ESG ratings across different rating providers. As mentioned by Berg et al. (2019), differences in ratings result from the choice of methodology and subjective interpretation. The absence of a reporting standard and a consensus among providers on a rating methodology remains an issue in the academic literature using ESG scores. Thus, highlighting a prominent problem in existing literature, namely, that the choice of rating provider can have large implications on results. Halbritter and Dorfleitner (2015) highlight the implications of ESG rating inconsistency by finding that the relationship between ESG and stock performance depends on the ESG rating provider.

Furthermore, Drempetic et al. (2019) results indicate that there is a size bias in the measurement of ESG ratings in the Thomson Reuters ASSET4 database. The results show significant positive correlation between firm size, available resources for providing ESG data and the availability on ESG data on sustainability performance. The authors argue that the ESG rating process favours larger firms with more resources.

3 Hypothesis

The thesis examines the relationship between ESG performance and stock performance. Based on the literature we present the following hypothesis:

3.1 Hypothesis

> *"Companies with high ESG performance have greater stock performance than companies with low ESG performance."*

Due to increased focus on the social responsibility of the firm by investors and society, we expect companies with high ESG performance to perform better than companies with low ESG performance. Most empirical evidence indicates a positive relationship between ESG performance and stock performance. The majority research is conducted on the US markets, we expect the relationship to still be prominent for Scandinavian markets.

As well as the two overall ESG metrics, ESG performance includes performance in the three ESG dimensions: environmental, social, and governance. The individual dimensions allow us to examine the differences and compare our results to the existing literature. Furthermore, analyzing the individual dimensions can improve the understanding of the overall ESG metrics and how the individual dimensions impact them.

4 Data

In this section, we will go through the sampling process and describe our different data providers and metrics.

4.1 Data Sources

The data was collected from two sources: Refinitiv Datastream and Kenneth R. French's data library. Refinitiv provides the GICS industry classification[6], monthly total returns, yearly ESG data, and yearly market capitalization. Furthermore, the Fama-French (Fama and French, 1993) and Carhart (Carhart, 1997) factors for European countries are obtained from Kenneth R. French's data library (French, 2021).

4.2 ESG-Data

Refinitiv is one of the world's largest providers of financial data, with historical data dating back to early 1900s (Refinitiv, 2021). The database contains information on company fundamentals, equities, bonds, commodities, and various other assets. All equity data are sourced directly from exchanges. Additionally, Refinitiv is one of the major ESG-rating providers, with a total coverage of more than 9000 companies, accounting for 80% of the global market capitalization. By assessing ESG scores using only publicly available information, Refinitiv preserves their objectivity. Refinitiv produces two overall ESG metrics to measure a firm's total environmental, social and governance performance. The two metrics are the ESG score and the ESGC score. The ESG score assesses the company's ESG performance, whereas the ESGC score combines the ESG score with ESG controversies. By including the controversy score in the assessment the ESGC score seeks to provide a comprehensive assessment of the company's long-term sustainability impact and conduct (Refinitiv, 2021).

[6]Global Industry Classification Standard - Developed by MSCI and Standard & Poor in 1999 (MSCI, 2021)

4.2.1 ESG Score

The scoring process starts with more than 500 different data points per company, of which 186 data points are chosen and used for scoring. The selected data points are the most relevant and comparable within each industry. Subsequently, the selected data points are divided into ten different categories, which are distributed across the three different ESG pillars.

Table 4.1: ESG pillars with corresponding pillar categories

Environmental	**Social**	**Governance**
Emission	Community	CSR Strategy
Innovation	Human rights	Management
Resource use	Product responsibility	Shareholders
	Workforce	

Refinitiv calculates the scores in the ten different categories using percentile rank scoring[7]. The categories are weighted differently within each pillar, and the pillar scores are the relative sum of the category weights. The score of each category is calculated by divided each category into 6-10 sub-categories[8]. Each sub-category is measured by several data points, either numeric or Boolean[9]. The environmental-, social- and governance pillar score are the evaluation of the firm's performance on the corresponding pillar. After the pillar scores are calculated[10], the overall ESG score is computed by multiplying the weights for each pillar with their respective score. The weight of the environmental and social pillars varies across different sectors, while the governance pillar remains constant across all sectors. The scoring process in summed up in the figure below, starting from the bottom.

[7]Percentile rank refers to the percentage of scores that are equal to or less than a given score. Since the score is based on rank, it will be unaffected by outliers.

[8]Presented in the appendix

[9]A number with two possible values, true(1) or false(0)

[10]Matrix for calculating pillar scores are presented in the appendix

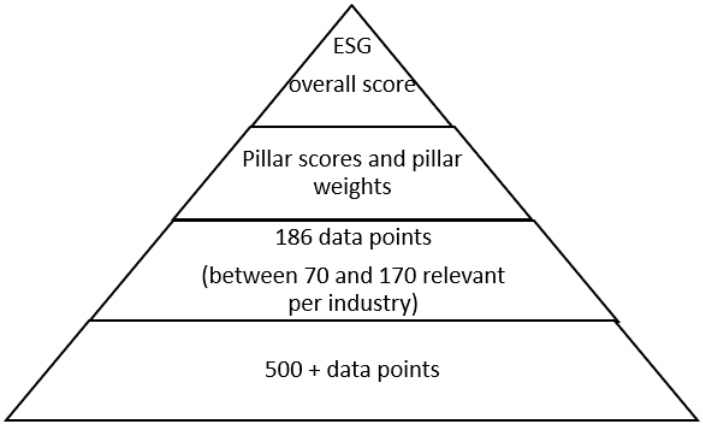

Figure 4.1: Summary of calculation of ESG score

4.2.2 ESGC - Controversies score calculation

The ESGC score is a combination of the ESG Score and the ESG Controversies score. The Controversies score is based on 23 controversy topics[11] and reflects the corresponding fiscal year's score. No controversies are double counted, and companies with no Controversies obtains a score of 100. As large-cap companies suffer more media attention than small-cap companies, the size of the market capitalization is considered by applying severity weights. The ESGC score is calculated as the average of the ESG score and the Controversies score when there are controversies during the latest fiscal year. If the Controversies score is larger than the ESG score, the ESGC will be equal to the ESG score (Refinitiv, 2021).

Table 4.2: ESGC calculation example

Scenario	**ESG Controversies Score**	**ESG Score**	**ESGC Score**
If the Controversies score is >= ESG score, then ESG score = ESGC score	100	70	70
If the Controversies score is <ESG score, then ESGC score = average of ESG and controversies score	50	70	60

4.2.3 Market Capitalization

Refinitiv calculates the market capitalization by multiplying the number of shares outstanding with the share price. The market capitalization is calculated on the rebalancing date of the portfolios. All market capitalizations are extracted in euros, allowing no need

[11]The 23 Controversies topics are listed in the appendix

for manual exchange-rate conversion. In which reduces the potential noise caused by the manual conversion in our results.

4.3 Sample Selection

To test our hypothesis, we constructed a sample consisting of Scandinavian companies. The Scandinavian sample consists of companies listed on Norwegian, Swedish, Danish and Finnish stock exchanges. Similar studies have been conducted on the American and European stock markets, but not for Scandinavian countries separately. By using a sample of Scandinavian companies, we contribute with research on the aforementioned relationship in a new region.

4.3.1 Screening

To start with, we had 1763 listed and unlisted Scandinavian companies. After filtering for listed companies, we were left with 1193 companies. Furthermore, we filtered for companies with available ESG data and were left with 587. We then applied a minimum of five-year continuous ESG data criteria to filter for companies with long-term ESG data availability. A total of 140 companies had five years of continuous ESG data. Lastly, we increased the threshold to ten years to increase the robustness of the results, leaving us with 113 companies. The 113 companies represent a broad selection of sectors and the four Scandinavian countries, as presented in Table 4.3. The sampling criteria may have caused biases. The issue is further discussed in Section 4.6.

Table 4.3: Sector and country statistics

Table 4.3 presents the allocation of companies between both industries and countries. Companies are sorted by the GICS industry identifier, which is provided by MSCI.

Industry	**Share**	**Country**	**Number of companies**
Consumer Decretionary	6%	Denmark	25
Health Care	9%	Norway	20
Industrials	27%	Sweden	45
Financials	16%	Finland	23
Consumer Staples	5%		
Energy	10%		
Materials	12%		
IT	4%		
Real estate	4%		
Utilities	1%		
Communication Services	7%		
Sum	**100%**	**Sum**	**113**

4.3.2 Sample Descriptions

Table 4.4: Summary Statistics

Table 4.4 presents the summary statistics for our sample. The ESG scores are retrieved from Refinitv and market capitalizations are in million euros. We include the mean score (Mean), the smallest observation (Min), the largest observation (Max), median (Median) and standard deviation (StdDev) of a firm's ESG rating, ESGC rating, environmental rating, social rating and governance rating in addition to market capitalization.

Statistic	MCAP	ESG Score	ESGC Score	Environmental	Social	Governance
N	1130	1130	1130	1130	1130	1130
Mean	8,581.777	58.135	56.075	58.640	61.420	53.425
St. Dev.	11,782.570	18.618	18.089	24.549	21.469	22.165
Min	2.760	3.840	3.840	0.000	1.160	6.530
Median	4,231.260	60.230	57.790	63.240	66.300	54.210
Max	110,581.800	92.520	92.350	97.250	96.410	98.400

Table 4.4 presents the summary statistics of the sample. There is a large gap in market capitalization, ranging from 2.76 million euros to 110 billion euros. The ESGC Score descriptives show a lower mean and variation compared to the ESG Score. The effect is anticipated because the inclusion of controversies can only lower the score. The variation in the overall ESG scores is lower than in the individual pillar scores, which is natural as the overall scores are aggregations of the pillar scores. As for the pillar scores, the environmental pillar have the most significant variations, while social scores have the least.

4.3.3 Portfolio Selection

Firstly, the portfolio construction is based on dividing the companies into quintiles sorted on ESG performance. The following scores are used as a measure of ESG performance: ESG, ESGC, environmental, social, and governance-pillar scores. Secondly, we construct portfolios consisting of companies from the top and bottom quintiles. The top quintile companies enter the long portfolios, and the bottom quintile companies enter the short portfolios. The ESG scores are reported yearly on December 31st. Subsequently, the portfolio construction was lagged by one month to avoid look-ahead bias. Furthermore, the stocks were held for one year before rebalancing the portfolios based on changes in ESG scores. To adjust for considerable differences in market capitalization, we will construct both equal-weighted and value-weighted portfolios.

4.4 Variable descriptions

4.4.1 Dependent Variable

The monthly portfolio return of the zero-investment long-short portfolio is the dependent variable in our regression models. It is calculated by multiplying each stock's return by the corresponding weight in the portfolio. Change in stock price, as well as any relevant dividends or stock splits, are factored into the total return. The formula for the total monthly return is as follows:

$$r_t = \frac{P_t}{P_{t-1}} - 1 \tag{4.1}$$

$Where:$

r_t = Total return at time t

P_t = Price adjusted for dividends and stock split

To compute weights for the value-weighted portfolios, we divide each stock's market capitalization by the total market capitalization of all stocks in the portfolio at the time of rebalancing, as denoted by Formula 4.2. The formula for equal- and value-weighted portfolios return is denoted by Formula 4.3.

$$w_{it} = \frac{MarketCap_{it}}{\sum_{n=1}^{N} MarketCap_{it}} \tag{4.2}$$

$$R_{pt} = \text{Total return of portfolio at time t}$$

$Where:$

$$w_{it} = \text{Weight of stock i at time t}$$

$$r_{it} = \text{Return of stock i at time t}$$

4.4.2 Fama French and Carhart factors

We extract the variables for the Fama French and Carhart models from Kenneth French's data library. The factors are reported monthly and are calculated for the Western European markets, corresponding to our sample. The factors are explained in the methodology section.

4.5 Limitations to data

We recognize that our data has limitations. The analysis is based on ESG scores, which automatically excludes companies that do not have an ESG rating. As a result, using ESG scores as a measure of ESG performance may result in selection bias. The concern is that the companies in our sample do not represent the ESG performance of companies in Scandinavia. Furthermore, the ten-year continuous data criteria may raise similar concerns. However, the companies in our sample vary in sector, firm size, and country, as well as ESG scores.

Furthermore, Refinitiv evaluates the ESG performance using publicly available data and penalizes companies that do not disclose information with lower ESG scores (Refinitiv, 2021). Thus, ensuring that companies that attempt to conceal information are punished rather than excluded from the rating. As a result, we argue that our sample is not influenced by selection bias.

Lastly, the portfolio returns do not take into account transaction costs. Not considering transaction costs produces the same results for portfolios with high and low turnover rates, which is not the reality.

5 Methodology

We construct portfolios with the top and bottom quintiles of ESG performers. This enables us to compare the performance of the top and bottom ESG score portfolios, using a long-short zero investment strategy with the Fama-French three-factor, Carhart four-factor and Fama French five-factor models. All three models are extensions of the CAPM framework (Sharpe, 1964; Mossin, 1966; Lintner, 1969), but due to lack of empirical success and the emergence of various risk factors, most applications of the CAPM model is invalidated Fama and French (2004). Therefore, we will not include the CAPM model in our research.

The methodology section is divided into two sections. First, we present the factor models we use to conduct research. Second, we present the underlying assumptions of our model.

5.1 Model specification

The factor models try to explain stock return variation by including factors that explain excess returns. The goal of the models to explain all the variations in stock prices with the risk factors. The factor coefficients and intercept were estimated using ordinary least squares regressions (OLS). The estimates are interpreted as differences between the long- and short portfolio. Thus, an insignificant coefficient indicates no difference in exposure to the corresponding risk factor. The intercept is the alpha of the investment strategy. The estimates and the R-squared may be less significantly, as the linear model estimates coefficient of differences and not a long only portfolio.

5.2 Fama-French Three-factor Model

The Fama-French three-factor model is an extension of the CAPM framework and includes the additional risk-factors “Small minus Big” (SML) and “High minus Low” (HML). Thus, accounting for the portfolio’s exposure to differences in size and value. The SML factor simulates a portfolio that is long on small market capitalization stocks, and short on large market capitalization stocks, which historically has generated positive risk-adjusted returns (Fama and French, 1993). The SML factor accounts for the portfolio’s exposure

to the outperformance of small-cap stocks by computing a small stock premium. The HML factor simulates a portfolio that is long high book-to-market value and short low book-to-market value companies. Thus, accounting for the historical overperformance of high book-to-market value companies by computing a value premium.

$$R_{i,t} - Rj, t = \alpha + \beta_{MKT} \times (MKT_t - R_{f,t}) + \beta_{SMB} \times \text{SMB}_t + \beta_{HML} \times HML_t + \varepsilon_{i,t} \quad (5.1)$$

Where:

$\alpha =$ The abnormal returns of the portfolio

$\beta_{MKT} =$ Exposure to the market factor (market beta)

$(MKT_t - R_{f,t}) =$ Excess return in the market at time t

$\beta_{SMB} =$ Exposure to the size factor

$SMB_t =$ The size factor at time t

$\beta_{HML} =$ Exposure to value

$HML_t =$ The value factor at time t

$\varepsilon_{i,t} =$ The error term

5.3 Carhart Four-Factor Model

In addition to the three former factors, Mark Carhart (1997) proposes a momentum factor accounting for the persistence of performance. This persistence comes from one year anomaly in returns for both high- and low-performance stocks found by Jegadeesh and Titman (1993). The momentum factor simulates a portfolio long the previous year's winning stocks, and short the previous year's losing stocks. Thus, accounting for the persistence of performance by computing a momentum premium for the portfolio.

$$R_t = \alpha + \beta_{MKT} \times (MKT_t - R_{f,t}) + \beta_{SMB} \times SMB_t + \beta_{HML} \times HMLHML_t + \beta_{MOM} \times MOM_t + \varepsilon_{i,t} \quad (5.2)$$

$\beta_{MOM} =$ Exposure to momentum factor

$MOM_t =$ Momentum factor at time t

5.4 Fama-French Five-Factor Model

Later, Fama and French (2015) expanded the three-factor model by adding two more factors. Because the three-factor model fails to account for much of the variation in returns due to differences in profitability and investment, the two new factors are introduced (Fama and French, 2015). The two new factors added to the five-factor model are RMW and CMA. The RMW factor, which stands for "Robust minus Weak," computes the difference in return between diversified portfolios with companies with high profits and those with low profits. The RMW calculates the profitability premium and accounts for the portfolio's exposure to differences in profitability. The CMA factor, which stands for "Conservative minus Aggressive," computes the difference between two diversified portfolios, one of which includes only companies with low investments and the other with high investments. According to the empirical evidence presented by Fama and French (2015), companies with a high level of investment and, as a result, a higher expected growth in book equity, have a lower expected return. The CMA factor calculates the investment premium and accounts for the portfolio's exposure to different investment strategies. The five-factor model is also expanded by Carharts momentum factor (MOM). This model is referred to as five-factor model + momentum.

$$R_{i,t} = \alpha + \beta_{MKT} \times (MKT_t - R_{f,t}) + \beta_{SMB} \times SMB_t + \beta_{HML} \times HML_t + \beta_{RMW} \times RMW_t + \beta_{CMA} \times CMA_t \varepsilon_{i,t} \quad (5.3)$$

β_{RMW} = Exposure to profitability factor

RMW_t = The profitability factor at time t

β_{CMA} = Exposure to investment factor

CMA_t = The investment factor at time t

5.5 Fama-French Five-Factor Model + Momentum

This model adds the Carhart momentum factor to the Fama-French five-factor model.

$$R_{i,t} = \alpha + \beta_{MKT} \times (MKT_t - R_{f,t}) + \beta_{SMB} \times SMB_t + \beta_{HML} \times HML_t + \beta_{RMW} \times RMW_t + \beta_{CMA} \times CMA_t + \beta_{MOM} \times MOM_t + \varepsilon_{i,t} \quad (5.4)$$

All variables are described in previous sections.

5.6 Model testing

An important assumption for time-series inference is that the data are stationary. Non-stationary data can result in biased or inconsistent estimators (Woolridge, 2019). To justify our results, our dataset must satisfy the assumptions of homoscedasticity and no serial correlation (Woolridge, 2019). If the assumptions are not satisfied, the results from the OLS regressions might be spurious. The Breusch-Pagan test is used to check for heteroskedasticity. If heteroskedasticity is present, the OLS regressions are done with robust standard errors. The tables (A2.1-4) in the appendix present the results from the test and show that there is heteroscedasticity in several of our models. To deal with heteroscedasticity, we run our models with White's heteroscedasticity corrected standard errors (White, 1980). Additionally, we ran a Breusch-Godfrey to test for autocorrelation. As the results in tables (A2.5-9) in the appendix show, we do not have an issue with autocorrelation.

6 Results

In this section, we will present our regression models from the Fama French and Carhart models.

6.1 ESG Score

Table 6.1: ESG score portfolios regression output

Table 6.1 presents the results from The Fama French three-factor, five-factor, five-factor + momentum model and the Carhart four-factor model. The dependent variable is the monthly returns for the long-short zero-investment portfolio sorted on the ESG score. It presents the factor loadings and abnormal returns (in percentage) of both the equal-weighted (EW) and value-weighted (VW) zero-investment portfolios. The portfolio has a long position in the top quintile and a short position in the bottom quintile. MRKtRF is the market factor, SMB is the small-minus-big size factor, HML is the high-minus-low value factor, MOM is the momentum factor, RMW is the robust-minus-weak profitability factor, CMA is the conservative-minus-aggressive investment style factor.

	Dependent variable:							
	Portfolio Returns							
	FF3		Carhart		FF5		FF5 + mom	
	EW	VW	EW	VW	EW	VW	EW	VW
α	0.277	0.207	0.312	0.192	0.215	0.277	0.254	0.305
	(1.113)	(0.767)	(1.193)	(0.673)	(0.791)	(0.924)	(0.897)	(0.952)
MRKtRF	−0.142**	−0.204***	−0.150**	−0.201***	−0.148**	−0.141*	−0.152**	−0.144*
	(−2.444)	(−3.161)	(−2.403)	(−2.933)	(−2.008)	(−1.767)	(−2.020)	(−1.795)
SMB	−0.620***	−0.652***	−0.619***	−0.652***	−0.610***	−0.605***	−0.604***	−0.601***
	(−4.902)	(−4.169)	(−4.855)	(−4.137)	(−4.845)	(−3.627)	(−4.782)	(−3.625)
HML	−0.046	−0.117	−0.080	−0.103	0.186	−0.489*	0.125	−0.532
	(−0.529)	(−0.872)	(−0.808)	(−0.646)	(1.102)	(−1.704)	(0.623)	(−1.588)
MOM			−0.053	0.022			−0.058	−0.041
			(−0.707)	(0.230)			(−0.714)	(−0.417)
RMW					0.417*	−0.357	0.406*	−0.365
					(1.739)	(−0.982)	(1.682)	(−0.998)
CMA					0.017	0.575*	0.062	0.607*
					(0.061)	(1.876)	(0.215)	(1.848)
Observations	120	120	120	120	120	120	120	120
R^2	0.215	0.264	0.217	0.264	0.230	0.296	0.233	0.297
Adjusted R^2	0.194	0.245	0.190	0.239	0.196	0.265	0.192	0.260

*p<0.1; **p<0.05; ***p<0.01

The results show that the difference in abnormal returns between companies with a high ESG score and those with a low ESG score is insignificant. The neutral relationship holds true for all regressions with both weighting strategies. Further, it conflicts the hypothesis that high-ESG companies outperform low-ESG companies in terms of stock performance and suggesting the stock market correctly values ESG information.

In terms of systematic risk exposure, our findings show that the high ESG score portfolio has a lower beta than the low ESG score portfolio. The beta coefficient is statistically significant at the 10%, 5%, and 1% levels, depending on the model and weighting strategy. Robustness decreases when controlling for investment strategy and profitability, implying that the relationship is weaker when controlling for additional risk factors. The significant difference indicates a relationship between ESG scores and systematic risk. Supporting the supplementary hypothesis that ESG performance is a risk factor that impacts the expected return of the portfolio.

Furthermore, the SMB factor is significant in all regressions at a 1% level. The SMB factor indicates that the high and low portfolios have different exposures to size. The negative SMB factor indicates that the high portfolio is more exposed to firms with a high market capitalization than the low portfolio. The difference in size exposure could be explained by the size bias found by Drempetic et al. (2019). The size bias indicates that larger firms get a higher ESG rating, which is consistent with their relative representation in the high- and low ESG portfolios.

6.2 ESGC Score

Table 6.2: ESGC score portfolio regression output

Table 6.2 presents the results from The Fama French three-factor, five-factor, five-factor + momentum model and the Carhart four-factor model. The dependent variable is the monthly returns for the long-short zero-investment portfolio sorted on the ESGC score. It presents the factor loadings and abnormal returns (in percentage) of both the equal-weighted (EW) and value-weighted (VW) zero-investment portfolios. The portfolio has a long position in the top quintile and a short position in the bottom quintile. MRKtRF is the market factor, SMB is the small-minus-big size factor, HML is the high-minus-low value factor, MOM is the momentum factor, RMW is the robust-minus-weak profitability factor, CMA is the conservative-minus-aggressive investment style factor.

	Dependent variable:							
	Portfolio Returns							
	FF3		Carhart		FF5		FF5 + mom	
	EW	VW	EW	VW	EW	VW	EW	VW
α	0.188	0.125	0.211	0.142	0.137	0.092	0.169	0.137
	(0.712)	(0.499)	(0.763)	(0.531)	(0.481)	(0.339)	(0.574)	(0.469)
MRKtRF	−0.136*	−0.117	−0.141*	−0.121	−0.129*	−0.077	−0.132*	−0.081
	(−1.879)	(−1.237)	(−1.856)	(−1.237)	(−1.738)	(−0.928)	(−1.739)	(−0.960)
SMB	−0.483***	−0.209	−0.482***	−0.208	−0.463***	−0.142	−0.459***	−0.136
	(−3.072)	(−1.076)	(−3.057)	(−1.069)	(−2.788)	(−0.616)	(−2.748)	(−0.583)
HML	−0.170*	−0.136	−0.191*	−0.152	−0.036	−0.239	−0.085	−0.307
	(−1.670)	(−1.012)	(−1.732)	(−1.114)	(−0.193)	(−0.888)	(−0.408)	(−1.026)
MOM			−0.033	−0.025			−0.047	−0.065
			(−0.399)	(−0.257)			(−0.509)	(−0.577)
RMW					0.318	0.099	0.309	0.087
					(1.123)	(0.301)	(1.089)	(0.267)
CMA					0.120	0.398	0.157	0.448
					(0.402)	(0.984)	(0.492)	(1.026)
Observations	120	120	120	120	120	120	120	120
R^2	0.184	0.099	0.185	0.099	0.195	0.114	0.196	0.117
Adjusted R^2	0.163	0.076	0.156	0.068	0.160	0.075	0.154	0.070

*p<0.1; **p<0.05; ***p<0.01

After controlling for ESG controversies, we find no indication of higher abnormal returns for the top quintile of companies for both the equal-weighted and value-weighted portfolios. The results oppose our hypothesis of a positive relationship between the ESGC score and stock performance. Indicating the market is able to price the value of the ESGC risk premium correctly.

High-rated companies have a significantly lower exposure to systematic risk at a 10% level in the equal-weighted portfolios. Compared to the ESG score results, the the difference in systematic risk exposure is less prominent when accounting for ESG Controversies.

Only the SMB factor is signifacnt at a 1% level for equal-weighted portfolios. The factor indicates that top ESGC companies are significantly larger than the bottom companies, while the factor for the value-weighted portfolio is insignificant. The robustness in our models for all value-weighted portfolios disappears with portfolios based on ESGC scores, indicating no significant differences between the top and bottom companies for either of the two risk factors.

6.3 Environmental pillar score

Table 6.3: Environmental pillar portfolios regression output

Table 6.3 presents the results from The Fama French three-factor, five-factor, five-factor + momentum model and the Carhart four-factor model. The dependent variable is the monthly returns for the long-short zero-investment portfolio sorted on the ESG environmental score. It presents the factor loadings and abnormal returns (in percentage) of both the equal-weighted (EW) and value-weighted (VW) zero-investment portfolios. The portfolio has a long position in the top quintile and a short position in the bottom quintile. MRKtRF is the market factor, SMB is the small-minus-big size factor, HML is the high-minus-low value factor, MOM is the momentum factor, RMW is the robust-minus-weak profitability factor, CMA is the conservative-minus-aggressive investment style factor.

	Dependent variable:							
	Portfolio returns							
	FF3		Carhart		FF5		FF5 + mom	
	EW	VW	EW	VW	EW	VW	EW	VW
α	0.130	−0.206	0.192	−0.193	0.065	−0.249	0.156	−0.205
	(0.450)	(−0.759)	(0.612)	(−0.691)	(0.208)	(−0.789)	(0.474)	(−0.616)
MRKtRF	−0.056	−0.038	−0.070	−0.041	−0.019	0.012	−0.028	0.007
	(−0.713)	(−0.565)	(−0.822)	(−0.599)	(−0.199)	(0.154)	(−0.291)	(0.095)
SMB	−0.439***	−0.012	−0.437**	−0.012	−0.377**	0.075	−0.363**	0.082
	(−2.583)	(−0.075)	(−2.570)	(−0.072)	(−2.200)	(0.411)	(−2.111)	(0.436)
HML	0.137	0.157	0.079	0.144	0.152	0.006	0.010	−0.062
	(1.271)	(1.214)	(0.698)	(0.922)	(0.696)	(0.022)	(0.043)	(−0.194)
MOM			−0.090	−0.020			−0.134	−0.065
			(−0.931)	(−0.196)			(−1.333)	(−0.593)
RMW					0.315	0.089	0.289	0.077
					(1.042)	(0.231)	(0.954)	(0.192)
CMA					0.400	0.475	0.505	0.526*
					(1.205)	(1.555)	(1.511)	(1.731)
Observations	120	120	120	120	120	120	120	120
R^2	0.081	0.018	0.087	0.018	0.105	0.042	0.117	0.045
Adjusted R^2	0.058	−0.007	0.055	−0.016	0.065	−0.0005	0.070	−0.006
Note:								*p<0.1; **p<0.05; ***p<0.01

The results in terms of abnormal returns are consistent, with all alphas being insignificant. There is no difference between the high- and low portfolio in terms of abnormal returns, indicating a neutral relationship between environmental performance and stock performance. As a result, we were unable to provide evidence to support our main hypothesis.

Considering the systematic risk exposure, the high portfolio shows no difference in volatility compared to the low portfolio. The insignificant difference in beta is consistent across all regressions, indicating a neutral relationship between the ESG environmental score and systematic risk exposure. Subsequently, there is no evidence that the environmental performance impacts the perceived risk of the portfolios.

The regressions show that the portfolios based on environmental performance are similar in terms of alpha and risk factor exposure. Thus, indicating that they consist of stocks with similar exposure to systematic risk, value, profitability and investment strategy risk factors, except for the significant SMB factor in the equal-weighted portfolio. The insignificant risk factors implies that there are no difference risk factor exposure between companies with high environmental scores and those with low. The negative SMB factor suggest that the high portfolio has a greater exposure to large market capitalization stocks than the low portfolio.

6.4 Social pillar score

Table 6.4: Social pillar portfolios regression output

Table 6.4 presents the results from The Fama French three-factor, five-factor, five-factor + momentum model and the Carhart four-factor model. The dependent variable is the monthly returns for the long-short zero-investment portfolio sorted on the social pillar score. It presents the factor loadings and abnormal returns (in percentage) of both the equal-weighted (EW) and value-weighted (VW) zero-investment portfolios. The portfolio has a long position in the top quintile and a short position in the bottom quintile. MRKtRF is the market factor, SMB is the small-minus-big size factor, HML is the high-minus-low value factor, MOM is the momentum factor, RMW is the robust-minus-weak profitability factor, CMA is the conservative-minus-aggressive investment style factor.

	Dependent variable:							
	Portfolio returns							
	FF3		Carhart		FF5		FF5 + mom	
	EW	VW	EW	VW	EW	VW	EW	VW
α	0.235	−0.069	0.283	−0.073	0.226	0.062	0.279	0.080
	(0.902)	(−0.236)	(1.010)	(−0.225)	(0.824)	(0.207)	(0.954)	(0.243)
MRKtRF	−0.218***	−0.133	−0.229***	−0.132	−0.222***	−0.105	−0.227***	−0.107
	(−3.581)	(−1.557)	(−3.400)	(−1.459)	(−2.869)	(−1.238)	(−2.850)	(−1.233)
SMB	−0.476***	−0.599***	−0.474***	−0.599***	−0.480***	−0.608***	−0.472***	−0.606***
	(−3.416)	(−3.308)	(−3.374)	(−3.286)	(−3.287)	(−3.257)	(−3.224)	(−3.207)
HML	−0.358***	−0.489***	−0.403***	−0.485***	−0.263	−0.831***	−0.345	−0.858***
	(−3.771)	(−3.417)	(−3.707)	(−2.942)	(−1.226)	(−3.326)	(−1.424)	(−2.809)
MOM			−0.071	0.006			−0.078	−0.026
			(−0.794)	(0.055)			(−0.806)	(−0.218)
RMW					0.131	−0.592*	0.117	−0.596*
					(0.443)	(−1.842)	(0.393)	(−1.846)
CMA					−0.008	0.214	0.053	0.235
					(−0.025)	(0.654)	(0.153)	(0.635)
Observations	120	120	120	120	120	120	120	120
R^2	0.335	0.318	0.337	0.318	0.336	0.344	0.339	0.345
Adjusted R^2	0.317	0.301	0.314	0.295	0.306	0.316	0.304	0.310

Note: *p<0.1; **p<0.05; ***p<0.01

Based on social screens, there is no indication of a difference in alpha between the high- and low-rated companies. The insignificant alpha is consistent when controlling for all company-specific risk factors and weighting strategies. The results oppose the main hypothesis of a positive relationship.

The equal-weighted portfolios all indicate that top-rated companies are less exposed to systematic risk than bottom-rated companies based on social screens. Indicating that social performance impact the systematic risk of the portfolio. The results for the value-weighed portfolio indicate that there are no significant differences, indicating that difference in not prominent between the large market capitalization companies in the portfolios.

For the SMB factor, the results are significant at a 1% level in all our models, indicating that well-performing social companies are less exposed to small market capitalization firms. The HML factor is significant in all models, except for the equal-weighted portfolio in the five- and six-factor models. The results indicate that high-rated companies have lower book-to-market values than low-rated companies. The RMW-factor is also significant at a 10% level for the value-weighted portfolio in the five- and six-factor models.

6.5 Governance pillar score

Table 6.5: Governance pillar portfolios regression output

Table 6.5 presents the results from The Fama French three-factor, five-factor, five-factor + momentum model and the Carhart four-factor model. The dependent variable is the monthly returns for the long-short zero-investment portfolio sorted on the governance pillar score. It presents the factor loadings and abnormal returns (in percentage) of both the equal-weighted (EW) and value-weighted (VW) zero-investment portfolios. The portfolio has a long position in the top quintile and a short position in the bottom quintile. MRKtRF is the market factor, SMB is the small-minus-big size factor, HML is the high-minus-low value factor, MOM is the momentum factor, RMW is the robust-minus-weak profitability factor, CMA is the conservative-minus-aggressive investment style factor.

	Dependent variable:							
	Portfolio returns							
	FF3		Carhart		FF5		FF5 + mom	
	EW	VW	EW	VW	EW	VW	EW	VW
α	0.381*	−0.132	0.373*	−0.108	0.369	−0.149	0.357	−0.124
	(1.705)	(−0.508)	(1.665)	(−0.407)	(1.538)	(−0.555)	(1.483)	(−0.451)
MRKtRF	−0.071	0.052	−0.069	0.046	−0.083	0.047	−0.081	0.044
	(−1.243)	(0.687)	(−1.130)	(0.598)	(−1.372)	(0.601)	(−1.308)	(0.558)
SMB	−0.428***	−0.262	−0.428***	−0.262	−0.440***	−0.263	−0.442***	−0.259
	(−3.206)	(−1.493)	(−3.165)	(−1.476)	(−2.909)	(−1.433)	(−2.856)	(−1.387)
HML	0.145	0.352***	0.152	0.329**	0.276	0.439*	0.295	0.401
	(1.342)	(3.072)	(1.078)	(2.498)	(1.127)	(1.937)	(1.005)	(1.567)
MOM			0.011	−0.035			0.018	−0.037
			(0.095)	(−0.362)			(0.140)	(−0.370)
RMW					0.153	0.134	0.156	0.127
					(0.507)	(0.459)	(0.508)	(0.422)
CMA					−0.078	−0.017	−0.091	0.011
					(−0.309)	(−0.048)	(−0.342)	(0.031)
Observations	120	120	120	120	120	120	120	120
R^2	0.118	0.165	0.118	0.166	0.121	0.167	0.121	0.168
Adjusted R^2	0.095	0.144	0.087	0.137	0.082	0.130	0.075	0.123
Note:								*p<0.1; **p<0.05; ***p<0.01

The equal-weighted governance-based portfolio exhibits a positive abnormal return at a 10% significance level in the Fama-French three-factor and the Carhart four factor model. However, the abnormal returns is not significant when accounting for the companies relative size in the portfolio using value-weights. The abnormal returns is not robust to the inclusion of the additional risk factors profitability and investment strategy. The lack of robustness suggests that some of the over-performance can be explained by the portfolios' difference in exposure to profitability and investment strategy.

Furthermore, the portfolios show no difference in terms of systematic risk. Similarly, there is also no difference in terms of exposure to the momentum, profitability or investment risk factors. The insignificant differences are consistent across all models and weighting strategies.

The HML factor is significant at a 1% level for value-weighted portfolio in the three-factor model. However, the robustness decreases when controlling for additional risk factors. The HML factor is significant at a 1% level for the value-weighted portfolio in the three-factor model. However, the robustness decreases when controlling for additional risk factors. The decreasing robustness indicates that other risk factors explain much of the variation associated with the HML factor in the three-factor model.

7 Discussion

The thesis investigates the relationship between ESG performance and stock performance in Scandinavia by applying the long-short strategy. The overall ESG scores indicate a neutral relationship with insignificant abnormal returns in all models. Hence, providing no support for the hypothesis of a positive relationship between the variables. When using the three pillars in ESG, the portfolio based on governance pillar scores creates positive abnormal returns. In contrast, the portfolios based on environmental and social pillar scores show no significant results.

7.1 Abnormal Returns

The portfolios sorted on ESG and ESGC scores show no significant abnormal returns, implying a neutral relationship between ESG performance and stock performance. The neutral relationship contrasts with the findings of Kempf and Osthoff (2007), Statman and Glushkov (2009) and Eccles et al. (2014). Their findings show significant abnormal returns for a long-short portfolio, which is long in the top-performing ESG companies and short in the bottom-performing ESG companies. There could be several reasons why the conflicting results occur. The mentioned studies examine the relationship between ESG performance and stock performance in other periods and markets. The positive relationship might not be present in Scandinavia between 2011 and 2020, indicating that the results are not transferable to other markets.

Alternatively, the effect could have disappeared since the initial sampling periods used by Kempf and Osthoff (2007), Statman and Glushkov (2009) and Eccles et al. (2014). Borgers et al. (2013) find that the outperformance of well-performing ESG companies disappears after the sampling period of Borgers et al. (2013), implying the market had earlier failed to price the intangible assets related to ESG performance. However, due to learning mechanisms in the stock market, the benefit of ESG performance has disappeared. Similarly, Halbritter and Dorfleitner (2015)) find a neutral relationship between ESG performance and stock performance in a later period, supporting the claim that a learning mechanism has taken place. Therefore, a possible explanation could be

that such learning effects have occurred, and the Scandinavian market correctly values the ESG performance.

Halbritter and Dorfleitner (2015) found evidence of significant differences in results when using different rating providers. As pointed out by Berg et al. (2019), the ESG ratings diverge significantly across different rating providers. Kempf and Osthoff (2007) and Statman and Glushkov (2009) use ratings from KLD, while Eccles et al. (2014) combines an ASSET 4 rating with a SAM rating. As ratings vary between rating providers, the choice of rating provider could impact the results. However, these implications are not tested and are beyond the current scope of our study.

Another possible reason for a neutral relationship is the complexity of measuring ESG performance and that ESG-ratings do not capture all the effects of ESG activities. A significant portion of the metrics for ESG scores are based on qualitative information, which may be hard to quantify and aggregate into a score. Due to a lack of standardized reporting of ESG information, Berg et al. (2019) conclude that the various ESG-rating providers produce different results. Further, rating providers have substantial differences in underlying scoring methodology and data gathering. Because of the substantial differences, the actual ESG performance is less likely to be reflected in the different ESG scores. Hence, investors who rely on ESG scores for screening allocate capital to other companies than intended.

Our results challenge the view of the shareholder theory, which claims businesses should refrain from any activities not maximizing value for the firm and leave philanthropy and ESG activities to individuals. The shareholder theory view would imply that ESG activities should negatively affect stock returns and increase the performance of companies that do not focus on ESG activities. Contrary to this, the stakeholder theory claims that businesses must create value for all stakeholders to succeed. According to stakeholder theory, the relationship between ESG performance and stock performance should be positive.

According to pieconomics by Edmans (2021), a neutral effect is explained if the difference in ESG performance between top and bottom companies is due to the impact of pie-growing and pie-shrinking activities. A possible explanation for the neutral effect is that ESG investments often have a net zero present value. While projects with a net zero NPV increase the pie, the shareholders only get the part that allows them to break even. Thus, growing the pie and resulting in increased ESG performance but neutral stock performance.

7.2 Systematic Risk

The high ESG score companies have lower systematic risk than the low ones, indicating that the ESG score negatively correlates with risk. The high ESG stocks' resilience might explain the difference in systematic risk during periods of crisis Lins et al. (2017) or a specific ESG risk factor Bénabou and Tirole (2010). The resilience argument would imply that the high-ESG score stocks demonstrate lower downside risk than their counterparts. Due to the trust between the firm's stakeholders through investments in social capital, firms with high ESG scores show more resilience through negative market shocks. The downside protection of social capital is consistent with stakeholder theory, suggesting that stakeholder relations impact the firm's success.

Interestingly, the difference in systematic risk is less robust and almost disappears when we account for ESG controversies. The difference implies that low-risk companies get more media attention than high-risk companies. The decrease in robustness is rather interesting as we expect the Controversies scores to affect the ESG score equally regarding systematic risk exposure. Subsequently, we would expect an effect, if any, to impact the relationship in the opposite way.

7.3 Exposure to size

The results show a significant difference in exposure to size between the ESG score portfolios. Companies in the high portfolio have a higher market capitalization than those in the low portfolio. The difference in exposure to size indicates that large companies embrace and engage in ESG activities. The reason may relate to the costs associated with

ESG activities. ESG activities are often costly projects and frequently require investments of a large scale to create utility. The fixed cost may be an issue for smaller companies that do not generate sufficient free cash flows or have established ESG activities and reporting infrastructure. Furthermore, the results of Drempetic et al. (2019) indicated that large firms are rewarded with higher scores in the ESG scoring process due to more resources and data availability. The difference in exposure to size exposure could indicate a size bias in Refinitiv's ESG scores.

The representation of large-cap companies in the high-portfolio could be due to increased scrutiny and pressure from stakeholders to engage in ESG activities. However, it does not appear to affect the stock performance of the companies. Krüger (2015) explores the effect of ESG media scrutiny, finding that media attention to ESG activities impacts the short-term stock returns of companies. Both attention to positive and negative ESG activities are associated with a negative stock reaction. However, he also finds a positive reaction to corporate initiatives of undoing previous bad behavior, meaning that the short-term negative effects could be offset by taking positive actions. Thus, the increased media attention could potentially balance out the short-term reduction in stock performance by better highlighting the positive action by the board. Ultimately, the long-term effects of increased ESG media attention on stock performance are unclear.

7.4 Environmental, Social and Governance

Our results show no indication of a relationship between ESG performance and stock performance when sorting on environmental and social pillars. However, the long-short investment strategy generates abnormal returns when sorted by governance pillar scores. The literature suggests a positive effect for all three pillars, but the majority of our results primarily indicate a neutral effect, failing to support the hypothesis of a positive relationship. The majority of the prior research finds a positive relationship between either of the pillars and stock performance. A possible reason our results contradict the results in the literature is the learning mechanisms in the market, as mentioned under the discussion of ESG score results. Furthermore, the former research is conducted in different markets, implying that relationships found using environmental and social screens are not transferable to the Scandinavian market.

The literature suggests a positive relationship between social performance and stock performance. Edmans (2011) argues that companies with high employee satisfaction outperform the market due to the market's inability to price intangible assets, and Servaes and Tamayo (2013) find positive relationships using customer relations. However, the social score includes several metrics, including employee satisfaction and customer relations. Our results do not indicate that the market fails to price the performance related to the social pillar, which accounts for employee satisfaction. Therefore, the results fail to support our hypothesis and the results of Servaes and Tamayo (2013) and Edmans (2011).

Our regressions show minor differences in risk exposure between the high and low portfolios sorted by environmental and governance scores. The lack of difference in risk exposure is interesting as they show substantial variations in ESG performance. The social pillar is similar to the ESG score portfolio, with significant exposure to size and lower systematic risk. An interpretation could be that the social pillar is the driver of differences in systematic risk in the ESG portfolio. In terms of size exposure, all three pillars show a similar relationship to the ESG score portfolio, which indicates that the largest companies are superior regardless of the pillar.

Based on the evidence of the pillar scores, the social categories seem to have importance for the systematic risk a company faces. This view is supported by the findings of Albuquerque et al. (2019), which finds that companies with good customer relations face significantly lower systematic risk due to stronger customer loyalty and higher product differentiation.

8 Conclusion

This thesis examined the relationship between ESG performance, measured by ESG score, and stock performance in Scandinavia. We find no significant results on abnormal returns using Refinitiv ESG scores, indicating that ESG performance has a neutral relationship with stock performance and implying that the stock market correctly values ESG performance. Hence, providing no support for the hypothesis of a positive relationship between the variables. Furthermore, our findings show that well-performing ESG companies are less exposed to systematic risk, implying that investors demand an extra risk premium for investing in companies with poor ESG performance.

We do not find a positive relationship between stock market performance and ESG performance when using ESG score, ESGC score, and environmental and social pillar score. However, we find evidence of a positive relationship between governance performance and stock performance. The difference in systematic risk is not prominent when controlling for controversies and is only visible when testing on the social pillar score. Further, the results show a large difference in exposure to size between the ESG score portfolios, finding that the high portfolio companies are larger than those in the low portfolio. The over-representation of large-cap companies with high ESG scores indicates that large companies embrace and engage in ESG activities.

In summary, we find no relationship between ESG performance proxied by ESG scores from Refinitiv and stock returns in the Scandinavian markets using the long-short strategy. In future research we would encourage using data from several different providers as ESG ratings to control for differences in rating methodologies. Additionally, it could further highlight potential differences between rating providers.

References

Albuquerque, R., Koskinen, Y., and Zhang, C. (2019). Corporate social responsibility and firm risk: Theory and empirical evidence. *Management Science*, 65(10):4451–4469.

Bebchuk, L. A., Cohen, A., and Wang, C. C. (2013). Learning and the disappearing association between governance and returns. *Journal of financial economics*, 108(2):323–348.

Bénabou, R. and Tirole, J. (2010). Individual and corporate social responsibility. *Economica*, 77(305):1–19.

Berg, F., Koelbel, J. F., and Rigobon, R. (2019). *Aggregate confusion: The divergence of ESG ratings*. MIT Sloan School of Management.

Borgers, A., Derwall, J., Koedijk, K., and Ter Horst, J. (2013). Stakeholder relations and stock returns: On errors in investors' expectations and learning. *Journal of Empirical Finance*, 22:159–175.

Business Roundtable (2019). *Statement on the Purpose of a Corporation*. Business Roundtable, Washington D.C, United States.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of finance*, 52(1):57–82.

Core, J. E., Guay, W. R., and Rusticus, T. O. (2006). Does weak governance cause weak stock returns? an examination of firm operating performance and investors' expectations. *the Journal of Finance*, 61(2):655–687.

Derwall, J., Guenster, N., Bauer, R., and Koedijk, K. (2005). The eco-efficiency premium puzzle. *Financial Analysts Journal*, 61(2):51–63.

Drempetic, S., Klein, C., and Zwergel, B. (2019). The influence of firm size on the esg score: Corporate sustainability ratings under review. *Journal of Business Ethics*, pages 1–28.

Eccles, R. G., Ioannou, I., and Serafeim, G. (2012). *The impact of a corporate culture of sustainability on corporate behavior and performance*, volume 17950. National Bureau of Economic Research Cambridge, MA.

Eccles, R. G., Ioannou, I., and Serafeim, G. (2014). The impact of corporate sustainability on organizational processes and performance. *Management Science*, 60(11):2835–2857.

Edmans, A. (2011). Does the stock market fully value intangibles? employee satisfaction and equity prices. *Journal of Financial economics*, 101(3):621–640.

Edmans, A. (2021). *Grow the Pie: How Great Companies Deliver Both Purpose and Profit–Updated and Revised*. Cambridge University Press.

Fama, E. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *The Journal of Financial Economics*, 33(1):3–56.

Fama, E. F. and French, K. R. (2004). The capital asset pricing model: Theory and evidence. *Journal of economic perspectives*, 18(3):25–46.

Fama, E. F. and French, K. R. (2015). A five-factor asset pricing model. *Journal of financial economics*, 116(1):1–22.

Freeman, R. E. (1984). *Strategic management: A stakeholder approach*. Boston : Pitman.

Freeman, R. E. and Phillips, R. A. (2002). Libertarian stakeholder theory. *Business Ethics Quarterly*, 12(3):331–349.

French, K. R. (2021). Kenneth R. French Data Library. Retrieved: November 23, 2021:from: https://bit.ly/3FeskqN.

Friedman, M. (1970). A Friedman doctrine-- The Social Responsibility Of Business Is to Increase Its Profits. [Online; accessed 3. Dec. 2021].

Gompers, P., Ishii, J., and Metrick, A. (2003). Corporate governance and equity prices. *The quarterly journal of economics*, 118(1):107–156.

GSIA (2021). Global Sustainable Investments Review 2020. *Global Sustainable Investments Alliance*, Retrieved: November 27, 2021:from: https://bit.ly/33JeGyb.

Halbritter, G. and Dorfleitner, G. (2015). The wages of social responsibility—where are they? a critical review of esg investing. *Review of Financial Economics*, 26:25–35.

Hong, H. and Kacperczyk, M. (2009). The price of sin: The effects of social norms on markets. *Journal of financial economics*, 93(1):15–36.

Hong, H., Kubik, J. D., and Scheinkman, J. A. (2012). Financial constraints on corporate goodness. Technical report, National Bureau of Economic Research.

Jegadeesh, N. and Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of finance*, 48(1):65–91.

Kempf, A. and Osthoff, P. (2007). The effect of socially responsible investing on portfolio performance. *European Financial Management*, 13(5):908–922.

Krüger, P. (2015). Corporate goodness and shareholder wealth. *Journal of financial economics*, 115(2):304–329.

Lins, K. V., Servaes, H., and Tamayo, A. (2017). Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis. *the Journal of Finance*, 72(4):1785–1824.

Lintner, J. (1969). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets: A reply. *The review of economics and statistics*, pages 222–224.

Markowitz, H. (1959). *Portfolio Selection: Eddicient Diversification of Investments*. Wiley, New York.

Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica: Journal of the econometric society*, pages 768–783.

MSCI (2021). GICS - The Global Industry Classification Standard. *MSCI*, Retrieved: December 04, 2021:from: https://www.msci.com/our–solutions/indexes/gics.

Refinitiv (2021). *Environmental, Social and Governance (ESG)scores from Refinitiv*. Refinitiv, London, United Kingdom.

Renneboog, L., Ter Horst, J., and Zhang, C. (2008). The price of ethics and stakeholder governance: The performance of socially responsible mutual funds. *Journal of corporate finance*, 14(3):302–322.

Servaes, H. and Tamayo, A. (2013). The impact of corporate social responsibility on firm value: The role of customer awareness. *Management science*, 59(5):1045–1061.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The journal of finance*, 19(3):425–442.

Statman, M. and Glushkov, D. (2009). The wages of social responsibility. *Financial Analysts Journal*, 65(4):33–46.

Taraldsen, L. (2021). ESG Duds Have Nowhere to Hide, World's Top Stock Owner Says. *Bloomberg*, Retrieved: December 18, 2021:from: https://bloom.bg/3e3WFMO.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48(4):817–838.

Woolridge, J. (2019). *Introductory Econometrics*. South-Western College Publishing.

Appendix

A1 ESG calculations

Table A1.1: Overview of categories and sub-categories within each pillar

Pillar	**Category**	**Sub-category**
Environmental	Emission	Emission
		Waste
		Biodiversity
		Environmental management systems
	Innovation	Product innovation
		Green revenues, R&D and CapEx
	Resource use	Water
		Energy
		Sustainable packaging
		Environmental supply chain
Governance	CSR Strategy	CSR Strategy
		ESG reporting and transparancy
	Management	Structure (independence, diversity, comittees)
		Compensation
	Shareholders	Shareholder rights
		Takeover defenses
Social	Community	Community
	Human rights	Human rights
	Product responsibility	Responsible marketing
		Product quality
		Data privacy
	Workforce	Diversity and inclusion
		Career development and training
		Working conditions
		Health and safety

Table A1.2: Example of pillar score calculation

Pillar	Category	Category score	Category weights	Sum of category weights	New category weights	Formula: New category weights	Pillar Scores	Formula: Pillar scores
Environmental	Emissions	0.98	0.15	0.44	0.35	(0.15/0.44)	0.94	(0.98*0.35)+
Environmental	Resource use	0.97	0.15		0.35	(0.15/0.44)		(0.98*0.35)+
Environmental	Innovation	0.85	0.13		0.29	(0.13/0.44)		(0.85*0.29)
Social	Community	0.89	0.09	0.31	0.28	(0.09/0.31)	0.94	(0.89*0.28)+
Social	Human rights	0.95	0.05		0.17	(0.05/0.31)		(0.95*0.17)+
Social	Product responsibility	0.92	0.04		0.13	(0.04/0.31)		(0.92*0.13)+
Social	Workforce	0.98	0.13		0.43	(0.13/0.31)		(0.98*0.43)
Governance	Shareholders	0.73	0.05	0.26	0.20	0.05/0.26)	0.32	(0.73*0.20)+
Governance	CSR strategy	0.34	0.03		0.13	(0.03/0.26)		(0.34*0.13)+
Governance	Management	0.19	0.17		0.67	(0.17/0.26)		(0.19*0.67)

Table A1.3: Controversy topics

Category	**Label**
Community	Anti-competition controversy
Community	Business ethics controversies
Community	Intellectual property controversies
Community	Critical country controversies
Community	Public health controversies
Community	Tax fraud controversies
Human rights	Child labor controversies
Human rights	Human rights controversies
Management	Management compensation controversion count
Product responsibility	Customer controversies
Product responsibility	Customer health and safety controversies
Product responsibility	Privacy controversies
Product responsibility	Product access controversies
Product responsibility	Responsible marketing controversies
Product responsibility	Responsible R&D controversies
Resource use	Environmental controversies
Shareholders	Accounting controversies count
Shareholders	Insider dealings controversies
Shareholders	Shareholder rights controversies
Workforce	Diversity and opportunity controversies
Workforce	Employee healt and satisfaction controversies
Workforce	Wages or conditions controversies
Workforce	Strikes

A2 Model Testing

Table A2.1: Breuch-Pagan test

This table presents the Breuch-Pagan test for hetereoscedasticity. It covers the portfolios we have constructed based on the different ESG measures, both equal- and value-weighted. H0 for Breuch-Pagans test is that there is Homoscedasticity present in our models. Hence if there is a P-value above 0.05, we reject the hypothesis that Homoscedasticity is present. If the p-values are below 0.05, we must run the regressions with robust standard errors. The numbers are F-stat and P-values in the paranthesis

Heteroscedasticy										
	ESG Score		**ESGC**		**Environmental**		**Social**		**Governance**	
Models	EW	VW	EW	VW	EW	VW	EW	VW	EW	VW
Three-factor	0.329	7.470	0.107	15.648	1.427	2.709	0.220	3.532	3.475	2.534
	(0.954)	(0.058)	(0.990)	(0.001)	(0.699)	(0.438)	(0.974)	(0.317)	(0.324)	(0.469)
Carhart	0.803	8.199	0.198	15.667	1.912	2.835	0.853	7.208	6.729	2.894
	(0.938)	(0.085)	(0.995)	(0.004)	(0.751)	(0.585)	(0.931)	(0.125)	(0.150)	(0.575)
Five-factor	2.691	11.906	2.371	18.628	5.508	8.884	0.959	1.711	3.714	4.287
	(0.747)	(0.036)	(0.795)	(0.002)	(0.357)	(0.113)	(0.965)	(0.887)	(0.591)	(0.508)
Five-factor + momentum	3.000	14.226	2.334	18.392	6.054	9.017	1.329	4.555	8.059	4.907
	(0.808)	(0.027)	(0.886)	(0.005)	(0.417)	(0.172)	(0.970)	(0.602)	(0.233)	(0.555)

Table A2.2: Breuch-Godfrey test

This table presents the Breuch-Godfrey test for autocorrelation. It covers the constructed portfolios based on the different ESG measures, both equal- and value-weighted. H0 for the Breuch-Godfrey test is that there is autocorrelation present in our regression. Hence if there is a P-value above 0.05, we reject the hypothesis that autocorrelation is present. The models tests show that there are no autocorrelation in our models. The numbers are Chi-value and P-value in the paranthesis

Autocorrelation										
	ESG Score		**ESGC**		**Environmental**		**Social**		**Governance**	
Models	EW	VW	EW	VW	EW	VW	EW	VW	EW	VW
Three-factor	0.915	0.001	0.975	0.019	0.370	1.033	0.081	0.085	0.417	1.332
	(0.339)	(0.981)	(0.324)	(0.89)	(0.543)	(0.310)	(0.775)	(0.770)	(0.518)	(0.248)
Carhart	0.830	0.0003	0.922	0.013	0.173	1.021	0.145	0.091	0.441	1.521
	(0.362)	(0.986)	(0.337)	(0.910)	(0.678)	(0.312)	(0.703)	(0.762)	(0.506)	(0.217)
Five-factor	0.777	0.03	0.942	0.157	0.514	0.922	0.095	0.044	0.361	1.362
	(0.378)	(0.864)	(0.332)	(0..692)	(0.473)	(0.336)	(0.757)	(0.832)	(0.547)	(0.243)
Five-factor + momentum	0.714	0.021	0.897	0.103	0.343	0.972	0.168	0.032	0.384	1.512
	(0.398)	(0.884)	(0.344)	(0.748)	(0.558)	(0.324)	(0.681)	(0.857)	(0.535	(0.218)