Norwegian School of Economics
Bergen, Fall 2019

ESG Investments

Exploring the Impact of Sustainability on Financial Performance

Thea Aasheim Johannessen and Åsne Lovise Tveiterås

Supervisor: Associate Professor Francisco Santos

Master thesis, Economics and Business Administration

Major: Financial Economics

NORWEGIAN SCHOOL OF ECONOMICS

Acknowledgements

This thesis is written in collaboration between two students, majoring in Financial Economics. The thesis is part of the master's degree in Economics and Business Administration at the Norwegian School of Economics.

Our choice of topic was motivated by our joint interest in both finance and sustainability. Writing a master thesis on this topic therefore seemed like the perfect chance to develop a deeper understanding of something that interests us. Working with the thesis has been both challenging, exciting and educational. In addition, the fact that the topic is a relevant topic today has been very motivating.

First, we would like to thank our supervisor Francisco Santos for valuable feedback and advice, at all stages of the process. Secondly, we would like to thank colleagues and acquaintances for inspiring conversations on sustainable investing. Lastly, we thank our fellow students for being supportive and keeping our motivation up throughout the semester.

Norwegian School of Economics

Bergen, December 2019

Thea Aasheim Johannessen

Åsne Lovise Tveiterås

Abstract

In this thesis, we investigate the relationship between sustainability, measured by ESG score, and financial performance. The time period of the study is 2008-2018. We perform three different analyses. The first is a global analysis of all 11 years. In the second analysis, we split the sample into two sub-periods. Finally, we investigate the relationship for companies in the United States and Europe separately. We find that there is no indication of a positive relationship. Moreover, the results suggest that the relationship is more negative than neutral. This is especially evident globally for small companies, globally in the period 2009-2013, as well as for American companies. We also look at the relationship at the industry level and find that there are differences between industries. The findings in Health Care suggest that investing in sustainable companies destroys value. In Real Estate and Technology on the other hand, the results demonstrate a neutral relationship. Consequently, sustainable investments in these industries will neither create or destroy value.

Contents

List of Tables

1 Introduction

The last decade has been influenced by the word sustainability. It is used with regards to how an average person should live his or her life, in politics and in business. In a time of climate change and rapid technological changes, there seems to be growing agreement that sustainability is key to the future. The concept has also gained a strong foothold in the finance sector, which is considered to be vital in order to meet the Paris Agreement and Sustainable Development targets (UNEP Finance Initiative, 2019). The PRI, an investment initiative in partnership with UNEP Finance Initiative and the UN Global Compact, states: "We believe that an economically efficient, sustainable global financial system is a necessity for long-term value creation. Such a system will reward long-term responsible investment and benefit the environment and society as a whole" (Principles for Responsible Investments, 2018). This view is gaining support. Governments across the world are increasing their regulatory focus on incorporating sustainability considerations into investments. Simultaneously, investors are demanding better integration of sustainability into investment processes, and clients are showing greater interest in sustainable investing (Blackrock Investment Institute, 2019). The motivation for sustainable investing includes client demand, pursuing social or environmental benefits, fulfilling mission or values, minimizing risks and fulfilling the fiduciary duty, and lastly improving returns (Global Sustainable Investment Alliance, 2018).

Despite the financial effect of sustainability being researched for the last 40 years (Utgård, 2017), there is still not consensus as to whether sustainability generates higher returns. Academics have found evidence of the relationship being both positive, negative and neutral. Different concepts of sustainability have been used to measure the effect on financial performance over the years. Among these concepts are corporate social responsibility (CSR), socially responsible investing (SRI) and responsible investing (RI). The newest addition, and in our perception the most relevant measure today, is ESG. ESG measures how a company performs in terms of environmental, social and governance conditions. In this thesis, we perform our own analysis on the performance of sustainable versus non-sustainable companies, as measured by ESG score.

Our contribution to existing literature with this thesis is twofold. While previous research

has focused its attention towards the US or Europe, in this thesis we perform the analysis on a global data set, of 7 711 companies. The sample period is January 2008 to December 2018. Furthermore, there is little research on how the relationship between sustainability and financial performance in firms, differs between industries. This is also something we investigate in this thesis.

The aim of the thesis is to investigate whether a strategy of buying a portfolio of companies with high sustainability performance, and selling a portfolio of companies with low sustainability performance, generates abnormal returns. To measure sustainability we use the Thomson Reuters ESG score from Refinitiv's Datastream. The portfolios are constructed 1st of January every year, and held for one year. We deploy six different models to analyze the performance of the portfolios. These are the Capital Asset Pricing Model (CAPM), the Fama French three factor model with and without momentum, the Fama French five factor model with and without momentum, and finally the Fama French five factor model with momentum and the liquidity factor. The thesis consists of three analyses. For each analysis we perform one overall analysis and one industry level analysis. The general finding from this thesis is that the relationship between sustainability and financial performance is clearly non-positive, and more negative than neutral.

First, using the entire global sample, we find that small companies with low ESG scores outperform small companies with high ESG scores, by 0.4%-0.6% monthly. For larger companies, there is no significant difference between companies with high and low ESG scores. This is an interesting finding as it implies that firm characteristics affect the direction of the relationship between ESG and return.

Next, we split the sample in two sub-periods. The motivation behind this is that attention to ESG has increased dramatically in recent years, and we therefore have reason to believe that the relationship might have changed over our sample period. This is also what our results indicate, as there is less evidence of a negative relationship between ESG and financial performance in the second period. In the first period however, low ESG performing companies generate between 0.4% and 1.6% higher monthly returns, than high ESG performing companies.

Lastly, we extract companies from the United States and Europe into two separate data sets. The motivation behind this is to investigate whether two comparable regions display

the same relationship between ESG and financial performance. In addition, the US and Europe are among the largest regions measured in value of sustainable investing assets, and the two most frequently researched regions (Global Sustainable Investment Alliance, 2018). Interestingly, our results suggest that the relationship is negative in the United States and neutral in Europe. In the US, low ESG companies outperform high, by 0.3%-0.5% monthly, whereas there is no significant difference in performance in Europe.

Furthermore, we evaluate the performance of the strategy for 11 industries. The analysis is performed on industry level for both the global sample, the two sub-periods, and the US and Europe. For an investor seeking to apply our strategy, knowledge of how it performs across different industries is valuable. Existing literature has devoted little attention to this topic, but one example of empirical evidence is the work of Baired et al. (2012). They find that the relationship between sustainability and financial performance differs between industries. This is consistent with our results. In this thesis we find that Health Care stands out with convincing evidence of a negative relationship. Low ESG performing Health Care companies outperform high, by 0.6%-1.7%. In Real Estate and Technology the findings suggest a neutral relationship. For the remaining eight industries the relationship moves between being negative and neutral.

As mentioned, there is not yet consensus in existing literature as to whether the relationship between sustainability and financial performance is positive, negative or neutral. To provide our thesis with an empirical context, we will now present some of the existing literature. According to instrumental stakeholder theory, attention to corporate social performance (CSP) improves stakeholder relationships and results in better performance (Orlitzky et al., 2003). Further, the slack resource theory states that high levels of financial performance may result in slack resources that provide the opportunity to invest in corporate social responsibility and responsiveness (Orlitzky et al., 2003; Waddock and Graves, 1997). Eccles et al.(2014) performed a similar analysis as the one in this thesis, where they compared the performance of high versus low sustainability companies in the US, over 18 years. Contrary to our results, they found that the high rated companies outperformed the low. Other examples of academics finding a positive relationship are Orlitzky et al. 2003, Friede et al. (2015), Lean et al. (2015), Kemp and Osthoff (2007) and Verheyden et al. (2016).

Other academics argue that sustainable and responsible business practices impact financial performance negatively. They argue that companies that act responsibly are incurring costs that might otherwise be avoided, or that should be paid by others. This creates a competitive disadvantage (Waddock and Graves, 1997). In this thesis we find that globally, small firms with low ESG score outperform small firms with high ESG score. Further, we find that the relationship is negative in the US for the entire period, and globally during the time period 2009-2013. These findings of a negative relationship are in line with the works of Renneboog et al. (2008) and Chang et al. (2012).

There are also arguments in favour of a neutral, or non-existing, relationship between sustainability and financial performance. According to Waddock and Graves (1997), there are so many intervening variables between social and financial performance, that there is no reason to expect a relationship to exist, expect by chance. Lastly, one should be aware of the possibility of a publication bias. That is, papers finding a positive relationship between CSP and CFP might be more likely to get published than papers finding a negative relationship (Utgård, 2017). The results in this thesis indicate that among large companies globally, the relationship is neutral. Further, we find no significant difference in return between sustainable and non-sustainable companies in Europe, and little difference in the period 2014-2018. This is consistent with the findings of Revelli and Viviani (2015) and Hamilton and Statman (1993), who both find that there is no cost or benefit to investing responsibly. Halbritter and Dorfleitner (2015) also strongly question whether there is a relationship between ESG and financial performance.

An outline of the thesis is as follows. Chapter 2 presents the data. In section 3, we present the results from the empirical analysis. Finally, we summarize our results and draw a conclusion in chapter 4.

2 Data

In this chapter, we explain the process of collecting the data and making it ready for the empirical analysis. When collecting the data we have used three different sources: Thomson Reuters Datastream through Refinitiv; Kenneth R. French's data library[1] and Lubos Pastor's Research[2]. Datastream is a global financial and macroeconomic database, which covers 175 countries and more than 35 million instruments or indicators (Refinitiv, nd). It is also one of the most comprehensive ESG databases in the industry (Refinitiv, 2019). From Datastream we retrieved yearly ESG scores, monthly returns, market values and market-to-book values, and industry information. For the different factors used in the regression models, we use Kenneth R. French's data library and Lubos Pastor's Research.

In the following, we provide a presentation of the ESG score used in this thesis. Then we describe our sample and the screening process. Finally, we explain the portfolio construction and introduce the dependent variable and the risk factors used in the analysis.

2.1 ESG Score

ESG investing is the consideration of environmental, social and governance factors alongside financial factors in the investment decision-making process (MSCI, 2019a; The Global Compact, 2004). ESG first emerged as a term in 2005, and builds on the concepts of corporate social responsibility (CSR), social responsible investing (SRI) and responsible investing (RI). However, an important difference from the former measures is that ESG investing is based on the assumption that ESG factors have financial relevance. Therefore, the argument to incorporate ESG in the investment strategy is a matter of return in addition to a matter of ethical and moral criteria (Kell, 2018).

We retrieved yearly ESG scores from Thomson Reuters Datastream. The Thomson Reuters ESG database in Datastream contains ESG scores for more than 8 000 companies worldwide, which makes it one of the biggest and most comprehensive ESG databases.

[1]Kenneth R. French Data Library: https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

[2]Lubos Pastor's Research: https://faculty.chicagobooth.edu/lubos.pastor/research/liq_data_1962_2018.txt

Thomson Reuters only uses publicly available information when assessing ESG performance, which preserves their objectivity. Furthermore, they evaluate companies on more key issues than any other competitor (Douglas et al., 2017). A further advantage to Thomson Reuters with regards to our work with this thesis is that it is not ESG exclusive but a financial market data provider. Hence, it enabled us to retrieve all the data needed for the portfolio construction from one source. Lastly, Thomson Reuters Datastream is available to us through NHH, and based on these different facts we decided to use this source. It is, however, worth mentioning that not including other ESG score providers could be considered a limitation to the thesis. Seeing that ESG scoring is still quite new, the same company could receive different scores from different agencies. Consequently, including ratings from other agencies would have made the study more robust.

The Thomson Reuters ESG score measures and calculates over 400 company key indicators, of which the 178 most comparable and relevant are used in the final scoring process. These 178 indicators are grouped into 10 categories, which again constitute the three pillars in the ESG score: Environmental, Social and Governance. The methodology used to calculate the scores is percentile rank scoring. This implies that a company's score depends on the performance of other companies. The relevant questions to ask are how many companies have the same or a worse value than the current one, and how many companies have a value at all. This method is used to calculate a score for all the different indicators, using formula (2.1):

$$\text{Score} = \frac{\text{No. of companies w/worse value} + \frac{\text{No. of companies w/same value incl.current}}{2}}{\text{No. of companies w/value}} \tag{2.1}$$

Qualitative indicators, or indicators with no publicly available information, are excluded. The benchmark to compare against differs between the categories. For the governance categories, the country of the respective company is used, because best governance practices normally are more consistent within countries. As for the environmental and social categories, industry groups are used. The reason behind this is that these kinds of issues tend to be more similar for companies within the same industry. To reach the total score, weights are assigned to each category based on the number of indicators in

each of them. Consequently, categories with more indicators receive a higher weight than categories with fewer indicators. The score ranges from 0-100, where 100 is the best possible score (Refinitiv, 2019; Thomson Reuters, 2017). The distributions of the weights can be shown in table 2.1.

Table 2.1: Categories and weights in the Thomson Reuters ESG score

Pillar	Category	Indicators in scoring	Weights
Environmental	Resource Use	19	11%
	Emissions	22	12%
	Innovation	20	11%
Social	Workforce	29	16%
	Human Rights	8	4.50%
	Community	14	8%
	Product Responsibility	12	7%
Governance	Management	34	19&
	Shareholders	12	7%
	CSR Strategy	8	4.50%
		178	100%

This table explains the construction of the Thomson Reuters ESG score. Each of the three pillars consist of different indicators that make up the pillar categories. Each category is assigned a weight that is used to calculate the total ESG score.

2.2 Sample Selection

In this thesis we are constructing portfolios based on ESG scores, to examine a strategy of buying companies with high ESG scores and selling companies with low ESG scores. We perform the analysis on an overall level, and on the industry level. To test this strategy we use the ASSET4 Universe list in Datastream. This list contains all the ESG rated companies in the Thomson Reuters database, on a global level. At the time of retrieval, the list consists of 8 126 companies. Despite the high number of companies, it is obvious that there are both sustainable and non-sustainable companies that are not rated on ESG. Therefore, it is not inconceivable that the results could be different if more companies were rated. Further, the list only consists of public companies, which could be considered a limitation of this thesis. However, private companies are less obligated to disclose ESG information, and we therefore find it reasonable to limit the study to public companies. In addition, since a condition of our analysis is that every company must have an ESG score, using this list simplified our data collection process, as an important part of the

data cleaning is done for us.

Our chosen time period is 1st of January 2008 to 31st of December 2018. There are several reasons behind the choice of this time period. First of all, we wanted to include as many years as possible, to obtain more observations. Secondly, ESG has become more relevant in recent years, and the number of companies with ESG scores therefore decreases the further we go back in time. We decided to use 2008 as our starting year, as this was when the global financial crisis happened. The financial crisis contributed to an increased focus on firms' accountability, ethical behavior, risk handling, and ability to manage many different stakeholders (Galbreath, 2013). This lead to growth in the incorporation of ESG by institutional and individual investors, and we therefore believe 2008 is a sensible starting year for our analysis. We include 2018 because we are interested in recent data, however 2019 is excluded since the reporting of the ESG scores in Datastream is not completed when we retrieve the data.

2.3 Screening

In this section, we will describe the process of cleaning the data and making it ready for the portfolio construction.

Since we are performing industry level analyses, we require that all companies in our data set are assigned to an industry. Datastream offers different industry classifications. We decided to use the equivalent to the GICS classification. Datastream provides the same industries, however some of the names differ slightly. The GICS, or Global Industry Classification Standard, was developed by MSCI and S&P Dow Jones in 1999. The aim was to offer an efficient investment tool to capture the breadth, depth and evolution of industry sectors. The standard is a four-tired, hierarchical industry classification system, with 11 sectors. In this thesis, we refer to the 11 sectors as industries. The industries are Basic Materials, Consumer Discretionary, Consumer Staples, Energy, Financials, Health Care, Industrials, Real Estate, Technology, Telecommunications and Utilities (MSCI, 2019b). A brief description of the industries can be found in the Appendix (A0.1). All companies without industry information were deleted.

With regards to the ESG score, we chose to use yearly and not monthly scores. The reason for this was to avoid losing observations. This is of particular importance in the industry

level analyses where the number of companies can be low. Because we construct portfolios in January every year, we require companies to have an ESG score in January. However, the month of reporting of the scores in Datastream is not the same for all companies, and therefore we would lose many companies by using monthly scores. Further, we observed that the ESG scores rarely change during a calendar year, reducing the potential loss in accuracy by using yearly scores instead of monthly. By using yearly scores, our only condition with respect to ESG score becomes that the companies must have an ESG score for one or more years of our period. To avoid survivorship bias, we do not require continuous years with a score. This means that we keep companies that lack scores in the middle of the time period. Companies that do not have a score for the entire period are removed.

We measure the performance of the portfolios every month, and therefore collect monthly returns. The portfolios are picked 1st of January every year. For a company to be picked we therefore require it to have a reported return in January. If a company only has available returns from for example March onwards, we delete the observations for this company for the current year. Contrary to this, a company that only has available returns for the first six months of the year will not be deleted. We interpret this scenario as a delisting, acquisition or bankruptcy, and give the stock a return of 0% for the remaining months of the year. By including companies that only exist or have ESG scores for part of the sample period, we reduce survivorship bias.

After the process of data cleaning, we end up with a sample of 7 711 companies. The number of companies across the different industries, and a presentation of the distribution of ESG scores across the sample, are found in table 2.2.

Table 2.2: Global distribution of companies, and ESG scores, across the industries in the period 2008-2018

	Number of Companies		ESG Score			
Industry	**Total**	**Yearly Average**	**Min**	**Mean**	**Median**	**Max**
Basic Materials	465	430	0	51.14	51.57	93.51
Consumer Discrretionary	1 176	698	7.75	49.65	48.65	94.78
Consumer Staples	438	267	7.14	51.97	52.80	97.43
Energy	519	311	11.46	50.88	49.86	93.77
Financials	1 237	690	9.28	49.89	46.96	95.50
Health Care	676	295	7.82	50.62	47.94	94.99
Industrials	1 216	735	0	49.52	49.68	97.90
Real Estate	541	294	12.78	49.29	47.71	95.83
Technology	643	330	7.51	51.38	50.26	92.74
Telecommunications	267	163	8.06	51.77	52.31	92.74
Utilities	313	197	9.65	51.99	52.49	91.96
Total	7 711	4 410	0	50.39	49.52	97.90

This table presents the total number of companies, and the yearly average number of companies, in each industry. Additionally, it provides the distribution of ESG scores within each industry. The ESG score runs from 0 to 100, where 100 is the best and highest possible score.

2.4 Portfolio Construction

To test the strategy of buying companies with high ESG scores and selling companies with low ESG scores, we construct portfolios based on ESG scores. We buy the portfolio consisting of the companies with high ESG performance and sell the portfolio containing the companies with low ESG performance. Finally, we perform an analysis of the difference in return between the high and low portfolio. In addition to ESG scores and industry information, we need the market value to construct the portfolios. Market value is retrieved monthly, in US dollars, from Datastream.

The portfolios are constructed using different thresholds of ESG scores. For example, we assign the companies with the 10% highest ESG scores to the high portfolio, and the companies with the 10% lowest ESG scores to the low portfolio. For the general analyses, we use thresholds of 10%, 25% and 40%, whereas for the industry analyses we use 25% and 40% thresholds. The reason that we in certain cases use higher thresholds for the industry level analyses, is that the number of companies in some industries is quite low. Consequently, using decile portfolios would not provide a sufficient amount of companies in the portfolios.

Table 2.3 presents some descriptive statistics of the average high and low global portfolios, using the 25% threshold, for every year of the analysis. As expected is the average market value for companies in the high portfolio significantly higher than for companies in the low portfolio. Large companies often have excess resources that can be used to engage in ESG issues, leading to a higher ESG score. Looking at the book-to-market value however, we observe more similar numbers. This implies that the companies in the high and low portfolios are valued quite similarly.

Table 2.3: Characteristics of the global high and low portfolios

Year	Portfolio	Book to Market	Market Value
2008	High	0.562	499 717
	Low	0.631	248 445
2009	High	0.868	549 374
	Low	0.897	206 981
2010	High	0.669	949 768
	Low	0.669	358 260
2011	High	0.675	1 265 820
	Low	0.664	434 029
2012	High	0.772	1 338 677
	Low	0.761	476 854
2013	High	0.705	1 547 416
	Low	0.791	679 544
2014	High	0.622	1 785 714
	Low	0.815	642 433
2015	High	0.684	1 767 752
	Low	0.674	437 037
2016	High	0.757	1 844 161
	Low	0.724	434 482
2017	High	0.664	1 956 254
	Low	0.627	327 876
2018	High	0.621	2 395 716
	Low	0.583	209 626

This table provides the average book-to-market value and market value for companies in the high and low portfolios, across the sample years. The book-to-market value is calculated by dividing one by the market-to-book value, which we retrieved from Datastream. The market value is displayed in millions of US dollars. The portfolios used in this illustration are the global 25% portfolios.

We pick portfolios in January every year and hold the portfolios for one year. In this thesis, we disregard transaction costs. This is a limitation of the thesis, as the investor in

reality would incur costs when rebalancing the portfolios. Return is measured monthly. We allow companies to disappear from the portfolios during a year, but we do not allow new companies to enter the portfolio during the year. By including companies that disappear during the year we avoid survivorship bias. Because we use yearly ESG scores, a company cannot move between the high and the low portfolio during a year. If a company disappears during the year, we redistribute the amount invested in this company, including its return, to the remaining companies in the portfolio.

We construct both equally- and value-weighted portfolios. In the equally-weighted portfolios, we assign equal weights to each stock. When constructing these portfolios we also use the previous month's market value to control for delisted companies. If the market value in the previous month was zero, the company will not be included in the portfolio for the current month. We include this condition in addition to removing companies without a return in January, as explained in 2.3 Screening, as an extra control. The invested amount and return will be equally distributed among the remaining companies.

In the value-weighted portfolios, we assign weights to the different stocks depending on the market value of their total outstanding shares. This implies that bigger companies will be assigned bigger weights than smaller companies. If a firm disappears in the value-weighted portfolios, the market value will equal zero, and the company will consequently be assigned a weight of 0%. Therefore, we do not have to use the same treatment as for the equally-weighted portfolios. The amount invested in the delisted company, and the return, will be distributed among the remaining companies, with respect to the company weights.

2.5 Dependent Variable

After constructing the portfolios, we are ready to perform the analysis. The dependent variable in our analysis is the excess return between the high and the low portfolio. That is, we measure the return of each portfolio every month, and subtract the return of the low portfolio from the high. To reach the excess return, we subtract the risk-free rate, collected from Kenneth R. French' Data library (French, K. R., 2019). The risk-free rate used by French, for both global, American and European data, is the US 1-month treasury bill rate.

The monthly stock returns are retrieved from Datastream, using the Total Return Index (RI), which assumes that dividends are reinvested. The return index is calculated as follows, where P_t is the price in period t:

$$RI_t = RI_{t-1} * \frac{P_t}{P_{t-1}} \tag{2.2}$$

The exception is on the ex-dividend date (D_t), when the index is given as:

$$RI_t = RI_{t-1} * \frac{P_t + D_t}{P_{t-1}} \tag{2.3}$$

Upon retrieval, Datastream converts the return index to percentage return, using the following formula:

$$r_t = \Delta RI_t = \frac{RI_t - RI_{t-1}}{RI_{t-1}} * 100 \tag{2.4}$$

To ensure that our results are valid we have performed a selection of statistical tests. The aim of these tests is to confirm that the assumptions for ordinary least squares (OLS) estimation on time series data are satisfied. We performed the Augmented Dickey-Fuller test for stationary, the Breuch-Godfrey test for autocorrelation and the Breuch-Pagan test for heteroskedasticity. In addition, we tested the linearity assumption by plotting the residuals against the independent variables. All tests support that the assumptions are satisfied, and can be provided upon request.

2.6 Risk Factors

To evaluate the performance of our strategy, we use the following asset pricing models: the Capital Asset Pricing Model (CAPM), the Fama French three factor model with and without momentum, the Fama French five factor model with and without momentum and the Fama French five factor model with momentum and the Pastor-Stambaugh liquidity risk factor. We retrieved the liquidity factor from Lubos Pastor's research[3] (Pastor, L.,

[3]Lubos Pastor's Research: https://faculty.chicagobooth.edu/lubos.pastor/research/liq_data_1962_2018.txt?fbclid=IwAR1Pfyz9EGC60Ep9pm3SQZ0ublBbwtumjYWrsMehNZRzPYCCdNmDxUgadvw

2018). The remaining factors were collected from Kenneth R. French's Data library[4] (French, F. R., 2019). Because we have a global data set we used the global factors, which are named "the developed market risk factors" in French's data library. For the geographic analysis of the United States and Europe, we use factors for North America and Europe respectively.

After screening the data and constructing high and low ESG portfolios, we are ready to evaluate the performance of our strategy. The results of the analysis are present in the next chapter, Empirical Analysis (3).

[4]Kenneth R. French Data Library: https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

3 Empirical Analysis

This section provides the results from the empirical analysis. We examine whether a strategy of buying high rated ESG companies and selling low rated ESG companies generates abnormal returns. We construct equally- and value-weighted portfolios based on ESG scores from Thomson Reuters Datastream, 1st of January every year. The portfolios are rebalanced on ESG after one year, and return is measured monthly. Our sample period is January 2008 to December 2018.

The chapter consists of three analyses, all with the same structure. The first analysis is the *universal analysis* in 3.1. The other two analyses will be compared with this. In the universal analysis, we use the entire global data set of 11 years. First, we test the strategy on the data set as a whole, before we divide it into 11 industries. We test the strategy on every industry, with two questions in mind: Do we find the same trend in the industries as in the overall analysis? And, are there differences between the industries that can explain what drives the results in the overall analysis?

In the second analysis, in 3.2, we split the data set in two time periods, and perform the same analysis on both periods. This analysis is referred to as the *periodic analysis*. Then, we perform industry level analyses for both periods. In this section, we are interested in examining if the results in the two time periods are consistent with the results in the universal analysis. A central question is whether there are differences between the two periods, both on the overall and industry level.

In the last analysis, we examine whether we find different results based on geography. We call this analysis the *geographic analysis* (3.3). By extracting all companies from the US into one data set and all companies from Europe into another data set, we investigate whether there are differences between the two regions, on an overall and industry level.

Six different models are used to estimate the abnormal return in this thesis. In the universal analysis, we apply all the models, whereas in the periodic and geographic analyses we reduce the number of models. All models are given a number, which they keep throughout the chapter. The models, and corresponding numbers, are:

1. Capital Asset Pricing Model (CAPM)
2. Fama French three factor model
3. Fama French three factor model with Momentum
4. Fama French five factor model
5. Fama French five factor model with Momentum
6. Fama French five factor model with Momentum and Liquidity

We now continue to present the results from the empirical analyses.

3.1 Universal Analysis

The first analysis is a general analysis, where we use the entire global data set, without dividing it into industries. We construct quartile portfolios (Panel A in 3.1), assigning the top 25% companies based on ESG score to the high portfolio, and the bottom 25% companies to the low portfolio. Similarly, we construct decile portfolios (Panel B in 3.1) using the top and bottom 10% ESG performing companies. In the following, we will present the results of the analysis, shown in table 3.1.

Table 3.1: Universal Analysis

Panel A: Quartile Portfolios

	Equally-Weighted Portfolios						Value-Weighted Portfolios					
	(1)	(2)	(3)	(4)	(5)	(6)	(1)	(2)	(3)	(4)	(5)	(6)
Mkt-rf	-0.03	-0.02	-0.03	-0.02	-0.02	-0.02	0.04	0.03	0.06	0.02	0.02	0.03
SMB		-0.08	-0.08	-0.09	-0.09	-0.08		0.19	0.21	0.14	0.17	0.14
HML		-0.03	-0.05	-0.05	-0.09	-0.09		0.16	0.27	0.01	0.20	0.19
WML			-0.03		-0.03	-0.03			0.15		0.17	0.19
RMW				-0.06	-0.05	-0.05				-0.33	-0.38	-0.40
CMA				0.03	0.07	0.06				0.07	-0.13	-0.22
LIQ						-0.01						-0.10
α	-0.39***	-0.40***	-0.39***	-0.38**	-0.38**	-0.38**	-0.53	-0.52	-0.56	-0.43	-0.42	-0.43
N	132	132	132	132	132	132	132	132	132	132	132	132
R2	0.01	0.01	0.01	0.01	0.02	0.02	0.00	0.01	0.02	0.01	0.03	0.03
Adj. R2	-0.00	-0.01	-0.02	-0.03	-0.03	-0.04	-0.01	-0.01	-0.01	-0.03	-0.02	-0.02

Panel B: Decile Portfolios

	Equally-Weighted Portfolios						Value-Weighted Portfolios					
	(1)	(2)	(3)	(4)	(5)	(6)	(1)	(2)	(3)	(4)	(5)	(6)
Mkt-rf	-0.07	-0.06	-0.07	-0.07	-0.08	-0-08	0.16*	0.11	0.07	-0.01	-0.01	-0.01
SMB		-0.16	-0.16	-0.20	-0.19	-0.19		0.66**	0.64**	0.41	0.42	0.45
HML		-0.12	-0.14	-0.14	-0.16	-0.16		0.55**	0.42	0.33	0.18	0.16
WML			-0.02		-0.02	-0-02			-0.17		-0.13	-0.11
RMW				-0.14	-0.13	-0.13				-0.98*	-0.93*	-0.96*
CMA				-0.03	-0.01	-0.01				-0.38	-0.22	-0.32
LIQ						-0.00						-0.11
α	-0.61**	-0.62**	-0.61**	-0.56**	-0.56**	-0.56**	-0.69	-0.63	-0.59	-0.26	-0.27	-0.28
N	132	132	132	132	132	132	132	132	132	132	132	132
R2	0.02	0.03	0.03	0.03	0.03	0.03	0.02	0.08	0.09	0.11	0.01	0.01
Adj. R2	0.01	0.00	-0.01	-0.01	-0.02	-0.03	0.01	0.06	0.06	0.07	0.07	0.07

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

This table presents the results from the universal analysis. Panel A provides the results from the quartile portfolios, while Panel B provides the results from the decile portfolios. Both panels include equally- and value-weighted portfolios. The dependent variable is the monthly excess return of the long-short zero-investment portfolio. The independent variables are interpreted as follows: Mkt-rf is the market risk premium, also referred to as the systematic risk. The SMB factor shows the portfolio 's exposure to small market cap stocks, while the HML factor is the exposure to high book-to-market stocks. The WML is referred to as the momentum factor, as it captures exposure to winners versus looser. RMW captures the exposure to companies with robust profitability, and CMA shows the portfolio's exposure to companies with a conservative investing strategy. The alpha is the intercept, and represents the abnormal return achieved by applying this strategy. The models is estimated using monthly data from 2008-2018.

The most interesting parameter in this analysis is the alpha (α) because it represents the abnormal return the investor realizes when applying the strategy of buying high ESG and selling low ESG. Looking at table 3.1 we observe significant alphas at the 1% and 5% level for all the models when the portfolios are equally-weighted. The value-weighted portfolios exhibit no significant alphas. Since small companies often generate more alpha than large companies, an explanation for this might be that the small companies are given smaller weights in the value-weighted portfolio. Consequently, the significance disappears. In the equally-weighted quartile portfolios, the abnormal return is approximately 0.4%. The negative sign indicates that the monthly returns of the low ESG portfolios on average are 0.4% higher than for the high portfolios. When using decile portfolios the abnormal

return in favor of the low portfolios is even higher, approximately 0.6%.

Looking at the risk factors we only find significance in the decile value-weighted portfolios at the 5% and 10% level. The positive exposure to SMB in model 2 and 3 indicates that small market cap companies have a risk premium, and that an investor following our strategy is likely to buy high ESG companies with small market capitalization. Further, the positive exposure to HML in model 2 implies that investors buy high ESG value companies over high ESG growth companies. For models 4, 5 and 6 we observe negative exposure to the RMW factor. This indicates that companies with weak profitability have a risk premium, and that the high portfolios comprise more companies with weak profitability than robust profitability. The fact that significance for all factors only appears in the value-weighted portfolios, might indicate that risk premium in the factors is higher for large companies than small.

However, in general there is very little significance in the risk factors. When a risk factor is insignificant, it means that there is no significant difference in exposure to this factor between the high and the low portfolios. An explanation for this could be that the companies in the high and low portfolios are quite similar. From table 2.3 we recall that the high and low portfolios exhibit similar book-to-market values, and insignificant risk factors are therefore to be expected. Furthermore, the insignificance contributes to explain the low adjusted R squared in all the regressions. These results are however not so surprising because we are looking at a zero-investment strategy, and hence analyzing differences. In such an analysis the adjusted R squared may be smaller, and the estimated risk factors less significant, than in an analysis of the long and short portfolios separately.

As robustness, we also analyzed portfolios of the top and bottom 40% ESG performing companies. The results of this analysis were similar to the already presented results. On the basis of this, and the fact that these portfolios are likely to consist of many firms with an average ESG score, we decided to not include these portfolios in the universal analysis. The results can be found in the Appendix (A0.2).

3.1.1 Universal Industry Analysis

Next, we repeat the analysis at the industry level. Results are presented in table 3.2. The portfolios are constructed on the top and bottom 25% and 40% ESG performing firms. The reason we use larger portfolios than in the general analysis is that the number of companies in each industry differs. Some industries are significantly smaller than others, resulting in very low numbers of companies if the 10% portfolios are used. This was illustrated in table 2.2, which provides the yearly average number of companies in each industry.

For the industry level analysis the main focus is the abnormal returns. We are interested in investigating whether we find the same trend in the abnormal return when looking at the industries separately, as in the general analysis. Furthermore, differences in abnormal returns between the industries might indicate which industries drive the results in the overall analysis. For simplicity, only the alphas from the regressions are presented in table 3.2. The risk factors will be briefly commented at the end of the section.

Table 3.2: Global industry specific alphas for the period 2008-2018

		Equally-Weighted Portfolios						Value-Weighted Portfolios					
Industry	Portfolio size	(1)	(2)	(3)	(4)	(5)	(6)	(1)	(2)	(3)	(4)	(5)	(6)
Basic	25%	-1.371	-1.409	-1.367	-1.410	-1.416	-1.430	-1.448**	-1.371**	-1.344**	-1.226*	-1.233*	-1.202*
Materials	40%	-1.442*	-1.427*	-1.399*	-1.579*	-1.589*	-1.603*	-0.717	-0.687	-0.679	-0.587	-0.587	-0.578
Consumer	25%	-0.395**	-0.401*	-0.388*	-0.294	-0.298	-0.295	-1.080*	-1.070*	-0.995	-0.584	-0.594	-0.586
Discretionary	40%	-0.382**	-0.384**	-0.380**	-0.297*	-0.299*	-0.293*	-1.013	-1.037	-1.110	-0.526	-0.506	-0.496
Consumer	25%	-0.334*	-0.338*	-0.344*	-0.347	-0.350	-0.350	-0.166	-0.160	-0.108	0.396	0.386	0.364
Staples	40%	-0.254	0.252	-0.267	-0.273	-0.273	-0.277	-0.140	-0.103	-0.066	0.358	0.351	0.317
Energy	25%	-0.083	-0.111	-0.149	-0.155	-0.151	-0.176	0.093	0.053	0.083	0.053	0.046	0.040
	40%	-0.029	-0.046	-0.080	-0.075	-0.075	-0.082	-0.123	-0.117	-0.175	-0.057	-0.068	0.074
Financials	25%	-0.265	-0.250	-0.203	-0.158	-0.165	-0.152	0.149	0.152	0.032	-0.220	-0.193	-0.195
	40%	-0.309*	-0.300	-0.254	-0.263	-0.271	-0.261	-0.188	-0.186	-0.310	-0.212	-0.179	-0.157
Health Care	25%	-0.479*	-0.475*	-0.453	-0.642**	-0.652**	-0.666**	-1.680**	-1.634*	-1.579*	-1.489	-1.498	-1.533*
	40%	-0.560**	-0.553**	-0.559**	0.704***	-0.707***	-0.718***	-1.907***	-1.857**	-1.767**	-1.650**	-1.667**	-1.687**
Industrials	25%	-0.265*	-0.257*	-0.249	-0.244	-0.244	-0.243	0.009	-0.034	0.085	-0.220	-0.242	-0.228
	40%	-0.330***	-0.229*	-0.228*	-0.235*	-0.235*	-0.317**	-0.609	-0.604	-0.614	-0.630	-0.628	-0.622
Real Estate	25%	-0.091	-0.093	-0.110	-0.110	-0.107	-0.111	-0.214	-0.238	-0.214	-0.226	-0.233	-0.215
	40%	-0.113	-0.110	-0.110	-0.088	-0.087	-0.082	-0.431	-0.455	-0.427	-0.310	-0.316	-0.303
Technology	25%	-0.031	-0.012	-0.028	-0.147	-0.144	-0.144	0.336	0.374	-0.355	0.053	0.058	0.058
	40%	-0.170	-0.139	-0.152	-0.230	-0.227	-0.232	0.153	0.213	0.189	-0.184	-0.179	-0.182
Telecom-	25%	-0.606**	-0.610**	-0.617**	-0.543**	-0.543**	-0.555**	0.056	0.055	0.010	0.225	0.234	0.229
munications	40%	-0.445**	-0.448**	-0.472**	-0.409*	-0.405*	-0.413*	0.770	0.076	0.655	1.287*	1.314*	1.328*
Utilities	25%	-0.312	-0.315	-0.345	-0.342	-0.337	-0.327	-0.711	-0.714	-0.791*	-0.949**	-0.936*	-0.934*
	40%	-0.156	-0.163	-0.181	-0.140	-0.137	-0.129	-0.388	-0.359	-0.458	-0.369	-0.349	-0.342

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

This table provides the alphas for all 11 industries from the CAPM (1), the Fama French three factor model without and with momentum (2 and 3), the Fama French five factor model without and with momentum (4 and 5), and the Fama French five factor model with momentum and liquidity (6). The left side of the table presents alphas from the equally-weighted portfolios, while the right side presents alphas from the value-weighted portfolios. For all industries we report alphas from both small (25% ESG threshold) and big (40% ESG threshold) portfolios.

An important difference between the overall analysis and the industry level analysis is that the portfolios are not longer well-diversified. Where the portfolios were composed of companies from all the different industries in the overall analysis, the portfolios now only contain companies from the same industry. An implication of this is that there is more idiosyncratic risk attached to the portfolios. This explains that the magnitudes of some alphas in table 3.2 are quite high.

Table 3.2 shows that the industries Energy, Real Estate and Technology all have alphas that are not significantly different from zero. This indicates that there is no significant difference in the performance between high and low ESG companies in these industries. The same applies to Financials, which only exhibits one significant alpha. A further implication of these results is that the mentioned industries are not likely to drive the negative abnormal returns found in the general analysis.

There are however several industries with significant alphas at all three levels, 1%, 5% and 10%. Furthermore, significant alphas appear for both the equally- and value-weighted portfolios, in contrast to the general analysis. Similarly to the general analysis, almost all the significant alphas are negative. In Basic Materials, we observe significant alphas for the big equally-weighted portfolios and the small value-weighted portfolios. Although the level of significance for the majority of the alphas is only 10%, which is considered to be weak statistical significance, the magnitude of the alphas is financially interesting. In this industry low ESG performing companies on average generate between 1.2% and 1.6% higher monthly returns than high ESG performing companies. It is likely that these high numbers are impacted by idiosyncratic risk, but even so, they indicate that low ESG companies outperform high.

For Consumer Discretionary we find negative significant alphas in the majority of the equally-weighted portfolios of about 0.3%. Considering the weak significance in the value-weighted portfolios, the findings suggest that the difference in performance between high and low companies is larger for small firms than for big.

The Health Care industry shows significant alphas for both portfolio sizes, with the majority being significant at the 1% or 5% level. In the equally-weighted analysis, our findings suggest that the returns of low companies on average are 0.6% higher than the returns of high companies. For the value-weighted portfolios, the difference is even larger,

with an abnormal return of approximately 1.7% in favor of the low portfolios. The increase in magnitude when we go from the equally-weighted to the value-weighted portfolios implies that the difference in performance between high and low companies is bigger for large companies than small.

In Industrials we only observe significance in the equally-weighted portfolios, implying that the difference in performance between the two portfolios is larger for small companies. Nonetheless, the weak level of significance makes us refrain from concluding in this industry.

The Telecommunications industry displays significant negative alphas for all equally-weighted portfolios, at an average of 0.5%. Interestingly, looking at the value-weighted portfolios all alphas are positive. Despite only three of these being significant, and only at the 10% level, this finding indicates that there is a significant difference in the relationship between ESG and return for small companies and larger companies.

For the last two industries, Consumer Staples and Utilities, the majority of the models exhibit insignificant alphas. This indicates that there is no significant difference between high and low ESG performing companies in these industries.

Next, we will comment on the risk factors, which are provided in table 3.3. This table summarizes the findings of the 11 industry analyses, for both the 25% and 40% portfolios, and shows the fraction of significant risk factors for each model. Factors of all three significance levels, 1%, 5% and 10%, are counted.

Table 3.3: Fraction of significant risk factors in the universal industry analysis

	Equally-Weighted Portfolios						**Value-Weighted Portfolios**					
	(1)	(2)	(3)	(4)	(5)	(6)	(1)	(2)	(3)	(4)	(5)	(6)
Mkt-rf	5/22	2/22	1/22	0/22	0/22	0/22	4/22	4/22	1/22	2/22	2/22	2/22
SMB		3/22	3/22	2/22	2/22	1/22		3/22	3/22	3/22	3/22	3/22
HML		2/22	2/22	8/22	6/22	6/22		2/22	3/22	2/22	3/22	3/22
WML			2/22		2/22	2/22			6/22		5/22	6/22
RMW				3/22	4/22	3/22				6/22	5/22	6/22
CMA				7/22	5/22	2/22				3/22	1/22	1/22
LIQ						1/22						0/22

This table summarizes the findings of the 11 industry analyses, for both the 25% and 40% portfolios, and shows the fraction of significant risk factors for each model. Factors of all three significance levels, 1%, 5% and 10%, are counted. For example, looking at the equally-weighted portfolios, we find that for model (2), the HML factor is significant twice.

In the overall universal analysis, we observed few significant risk factors, and only in the value-weighted decile portfolios. Consequently, we interpreted this as little difference in exposure to the risk factors between high and low ESG rated companies. Compared to the universal analysis, there is more significance in the risk factors when we analyze the industries separately. Further, there is more significance in the value-weighted portfolios than in the equally-weighted portfolios, which corresponds to the findings in the overall analysis. This pattern is especially evident for the WML and RMW factors. A likely explanation is that the difference in exposure to these two factors is bigger and more significant for larger companies than for smaller. Studying the CMA and HML factors, we observe the opposite, that there is more significance for the equally-weighted portfolios. This implies that the differences in exposure to the investment and value factors are larger for smaller companies than for bigger.

3.1.2 Summary Universal Analysis

The main finding from the universal analysis is that the majority of the results exhibit negative alphas. From the general analysis, we observed that the low portfolios outperformed the high by 0.4%-0.6%. However, significance only appeared for the equally-weighted portfolios, indicating that the difference in performance between high and low portfolios is larger for small firms than it is for big. The findings from the industry level analysis indicate the same. Furthermore, the strategy of buying high and selling low ESG stocks seems to generate different results across the industries. Although the alphas on average are negative in all the industries, not all industries exhibit enough significance, or consistency across the models, to conclude whether the relationship is negative or neutral. However, for Energy, Financials, Real Estate and Technology we observe convincing evidence of a neutral relationship. For Health Care on the other hand, the results indicate a negative relationship. In this industry the low portfolios outperforms the high by 0.6%-1.7% per month. Despite these results, we do not find it to be enough significance or consistency across the models to draw a general conclusion. We therefore find it interesting to continue the analysis, to see if the picture becomes clearer. In the next section, we split the sample into two sub-periods and apply the strategy to both periods.

3.2 Periodic Analysis

In this analysis, we split the sample into two time periods. The motivation behind this is to examine whether we find the same results in the two time periods as in the universal analysis, and if we can observe differences between the two periods. The time period we study in this thesis has been influenced by both a financial crisis and a period of great expansion in the financial markets. We can therefore not exclude that the strategy we analyze could have performed differently if the market conditions were other. By dividing the data set in two sub-periods, we take into account that the performance of the strategy can change over time. Further, the attention to ESG has increased dramatically in recent years. Both governments and customers are demanding more from companies when it comes to sustainability and social responsibility. We therefore expect that the number of companies scored on ESG has increased in recent years, and the results for the last period may therefore be different than for the more previous period.

The sample is split into two periods of five years each. The first period is 2009-2013, and the second period is 2014-2018. We exclude the year 2008 because this is when the financial crisis took place. The financial crisis was a dramatic event that affected companies in all industries all over the world. Since we are now performing the analysis on fewer years, hence obtaining fewer observations, it is not unlikely that this event affects our results. In this part of the analysis, we therefore see it as appropriate to exclude 2008.

As in the previous analysis, the main focus in this analysis is also the abnormal return, or alpha. Table 3.4 reports the alphas from equally- and value-weighted portfolios for the two time periods. We use the same portfolio sizes as in the universal analysis, deciles and quartiles. Further, we have reduced the number of regression models to three. These are the Fama French three factor model (2), the Fama French three factor model with momentum (3) and the Fama French five factor model (4). From the universal analysis, we observed that the alphas from model 1 tended to be similar to model 2, and the alphas from model 5 and 6 tended to be similar to model 4. After running all models for the periodic analysis we found the same trend. To increase the readability of the tables and make the analysis easier to follow, we therefore decided to only present the results from models 2, 3 and 4. In addition, our perception is that these models are the most commonly used in finance. Models 1, 5 and 6 can be provided upon request.

Table 3.4: Periodic Analysis

	Panel A: 2009 - 2013						Panel B: 2014 - 2018					
	I: Quartile Portfolios						I: Quartile Portfolios					
	Equally-weighted Portfolios			Value-weighted Portfolios			Equally-weighted Portfolios			Value-weighted Portfolios		
	(2)	(3)	(4)	(2)	(3)	(4)		(2)	(3)	(4)	(2)	(3)	(4)
Mkt-rf	0.009	0.019	-0.016	-0.001	0.022	-0.075	Mkt-rf	-0.110	0.181**	-0.133	0.228*	0.227*	0.190
SMB	-0.112	-0.093	-0.164	0.270	0.313	0.123	SMB	-0.043	-0.006	-0.024	0.291	0.292	0.288
HML	-0.004	0.012	-0.041	0.548	0.584	0.335	HML	-0.118	-0.415**	0.100	0.013	0.008	0.204
WML		0.036			0.083		WML		-0.365**			-0.006	
RMW			-0.237			-0.786	RMW			0.218			0.113
CMA			-0.004			0.017	CMA			-0.353			-0.449
α	-0.367**	-0.382**	-0.243	-1.308*	-1.342*	0.938	α	-0.470*	-0.374	-0.518*	-0.250	-0.248	-0.284
N	60	60	60	60	60	60	N	60	60	60	60	60	60
R2	0.015	0.028	0.044	0.039	0.043	0.056	R2	0.041	0.142	0.062	0.077	0.077	0.091
Adj. R2	-0.038	-0.043	-0.044	-0.013	-0.027	-0.031	Adj. R2	-0.010	0.079	-0.026	0.028	0.010	0.007
	II: Decile Portfolios						II: Decile Portfolios					
	Equally-weighted Portfolios			Value-weighted Portfolios			Equally-weighted Portfolios			Value-weighted Portfolios		
	(2)	(3)	(4)	(2)	(3)	(4)		(2)	(3)	(4)	(2)	(3)	(4)
Mkt-rf	-0.024	0.001	-0.067	0.207	0.123	-0.023	Mkt-rf	-0.221	-0.323*	-0.311*	0.220	0.253	0.110
SMB	-0.130	-0.085	-0.231	1.146**	0.991*	0.751	SMB	-0.130	-0.077	-0.172	0.436	0.419	0.189
HML	-0.028	0.010	-0.138	0.906*	0.776	0.868	HML	-0.316	-0.735**	0.241	0.058	0.195	0.108
WML		0.087			-0.300		WML		-0.516			0.169	
RMW			-0.477*			-1.470*	RMW			0.364			-0.551
CMA			0.056			-1.157*	CMA			-1.063			-0.691
α	-0.425*	-0.460*	-0.187	-1.606**	-1.485*	-0.812	α	-0.870*	-0.734	-0.953*	-0.379	-0.423	-0.278
N	60	60	60	60	60	60	N	60	60	60	60	60	60
R2	0.017	0.060	0.090	0.180	0.217	0.240	R2	0.057	0.113	0.099	0.045	0.049	0.073
Adj R2	-0.035	-0.009	0.005	0.136	0.160	0.169	Adj. R2	0.007	0.048	0.016	-0.006	-0.020	-0.013

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

This table presents the results from the periodic analysis. Panel A shows the results from the first period, 2009-2013, while Panel B shows the results from the second period, 2014-2018. For both periods we use quartile and decile portfolios, for both the equally- and value-weighted method. The models used in this analysis is the Fama French three factor model (2), the Fama French three factor model with momentum (3) and the Fama French five factor model (4).

From table 3.4 we observe that the alphas are negative in both time periods. This is consistent with the universal analysis and indicates that the low portfolios outperform the high in both time periods. There are however differences between the two periods. First, we notice that there are more significant alphas in the first period than in the second. In addition, the alphas in the first period are significant at the 5% or 10% level, whereas the alphas in the second period only are significant at the 10% level. Furthermore, the first period exhibits significant alphas in both the equally- and value-weighted portfolios, at approximately 0.4%-1.6%, while significance only appears in the equally-weighted portfolios in the second period, at 0.5%-1%. This indicates that the trend that low companies outperform high was stronger in the first period than in the second. In other words, the difference in performance between high and low firms is less today than it was 5 to 10 years ago. This result is not surprising because of the increased focus on ESG in

recent years. Both governments and customers are imposing requirements and demands on businesses to act more sustainable and responsible. Additionally, the most significant increase in the number of requirements for ESG reporting has taken place in the last 5 years (Climate Disclosure Standards Board, 2019). When more companies report on their ESG issues, the foundation for ESG scoring also increases, and more companies can be scored. This is also visible in our data, as the number of companies with ESG scores increases in all industries, from the first period to the second. This evidence is provided in table 3.5.

Table 3.5: Yearly average number of companies per industry in the two time periods

	2009 - 2013	**2014 - 2018**
Basic Materials	419	474
Consumer Discretionary	609	837
Consumer Staples	242	307
Energy	305	334
Financials	562	869
Health Care	207	408
Industrials	656	864
Real Estate	232	382
Technology	277	410
Telecommunications	150	185
Utilities	190	215
Total	3 849	5 285

This table provides the yearly average number of companies for each industry, in the two time periods. From the table we observe that the number of companies has increased in all industries.

Regarding risk factors, insignificance still dominates the results. Because we are using the same data set, the companies and their characteristics have not changed. Hence, the companies in the high and low portfolios are still quite similar, and insignificant risk factors are therefore as expected. There is however more significance when we split the sample, with the majority appearing in the second period.

3.2.1 Periodic Industry Analysis

Next, we will look into the results on industry level for the two time periods. These are provided in table 3.6. In this analysis, we only use the quartile portfolios. The reason we choose this over the 40% portfolios is that we want to ensure that the companies in the high and low portfolios have sufficiently different ESG scores. Additionally, there are enough companies at the industry level to use the quartile portfolios. We are interested in examining whether the results in the two periods are similar to the results in the universal industry analysis in 3.1.1, and whether there are differences within the industries between the two time periods.

Table 3.6: Industry specific alphas for the two time periods

	Panel A: 2009 - 2013						**Panel B: 2014 - 2018**						
	Equally-Weighted Portfolios			**Value-Weighted Portfolios**				**Equally-Weighted Portfolios**			**Value-Weighted Portfolios**		
	(2)	(3)	(4)	(2)	(3)	(4)		(2)	(3)	(4)	(2)	(3)	(4)
Basic Mat.	-0.404	-0.425	-0.184	-0.931	-0.899	-0.337	Basic Mat.	-2.797	-2.380	-3.106	-1.802	-1.566	-1.668
Con. D.	-0.692**	-0.688**	-0.265	-2.140*	-1.949	-1.463	Con. D.	-0.296	-0.253	-0.341	-0.169	-0.372	-0.141
Con. S.	-0.448	-0.495	-0.220	-0.770	-0.646	-0.849	Con. S.	-0.258	-0.145	-0.246	0.071	0.033	-0.006
Energy	-0.393	-0.505	0.059	0.614	0.642	0.432	Energy	0.459	0.562	0.220	-0.628	-0.600	-0.788
Financials	-0.206	-0.146	-0.196	-0.929	-1.061	-1.464*	Financials	-0.209	-0.162	-0.181	1.344*	1.355*	1.103
Health Care	0.058	0.053	0.062	-1.565	-1.596	-2.028	Health Care	-1.015*	-0.806*	-1.127**	-1.692	-1.299	-1.708
Industrials	-0.045	-0.030	-0.044	1.288*	1.502**	1.155	Industrials	-0.380*	-0.341	-0.352	-1.072	-1.036	-1.338*
Real Estate	-0.195	-0.236	-0.114	-0.171	-0.208	-0.178	Real Estate	-0.200	-0.180	-0.246	-0.739	-0.578	-0.664
Technology	-0.091	-0.061	-0.194	0.416	0.453	-0.145	Technology	0.054	0.106	-0.047	-0.289	-0.275	-0.530
Telecom.	-0.819***	-0.824**	-0.875**	-0.472	-0.663	-0.317	Telecom.	-0.549	-0.527	-0.419	0.539	0.811	0.616
Utilities	-0.360	-0.427	-0.344	-0.814	-0.987	-1.183	Utilities	-0.311	-0.259	-0.347	-0.789	-0.589	-0.779

Note: *p<0.1; **p<0.05; ***p<0.01
This table provides the alphas from the periodic industry level analysis. Panel A shows the alphas from the first period, 2009-2013, while Panel B shows the alphas for the second period, 2014-2018. For both periods we analyze equally- and value-weighted portfolios. All portfolios are constructed using a 25% ESG threshold. The applied models are the Fama French three factor model (2), the Fama French three factor model with momentum (3) and the Fama French five factor model (4).

For Consumer Staples and Utilities we observe no significance when splitting the sample. This is in line with the findings from the universal industry analysis and strengthens the perception of a neutral relationship in these industries. Basic Materials exhibited significance, though not strong, in the majority of the models in the universal industry analysis. However, when we split the sample into two periods, all significance disappears. For Consumer Discretionary and Telecommunications there is only significance in the first period when we split the sample. This indicates that the difference in performance between high and low ESG performing companies in these industries has diminished

during the sample period. For Telecommunications, we also observe some positive alphas, though not significant, in the second period. This is in line with the positive significant alphas we observed in the universal industry analysis.

The Health Care industry was the industry with the most significant alphas in the universal industry analysis. Therefore, an interesting observation is that after splitting the data, significant alphas only exist for the equally-weighted portfolios in the second period. This indicates that the low Health Care companies only started to outperform the high in the second period. In the first period, there was no significant difference in the return between high and low Health Care stocks.

Moving to Industrials, we recall finding significant alphas for the equally-weighted portfolios in the universal industry analysis. Now we observe two positive alphas in the first period, and two negative alphas in the second period. This indicates that high ESG companies outperformed low in the first period, while low outperformed high in the second period. However, the majorities of the alphas in both periods are only significant at the 10% level, as also was the case for the majority of the alphas in the universal industry analysis. The low significance from both analyses, and the mix of positive and negative alphas, makes it difficult to conclude that there is a significant difference in the performance between high and low companies in Industrials.

In the universal industry analysis, we concluded that there was no significant difference in return between high and low firms in four industries. For Energy, Real Estate and Technology this does not change when we split the sample. This reinforces the idea that there is no significant difference between high and low companies in these industries. For Financials however, we observe significant alphas in both periods when we split the data set. Moreover, the two significant alphas in the second period are positive, indicating that the high portfolios outperform the low. The level of significance is however only 10%. This, in combination with the negative alpha with a similar level of significance in period one, makes us doubt that there is a significant difference in return between high and low companies in the financial industry.

Continuing with the risk factors, presented in table 3.7, we observe that the industry level analysis reveals more significant risk factors than the overall analysis. This is consistent with the findings from the universal analysis. Further, there is more significance in period

1 than in period 2. In other words, high and low performing ESG companies are more similarly exposed to the risk factors today, than 5-10 years ago.

Table 3.7: Fraction of significant risk factors in the periodic industry analysis

	Panel A: 2009 - 2013						**Panel B: 2014 - 2018**						
	Equally-Weighted Portfolios			**Value-Weighted Portfolios**				**Equally-Weighted Portfolios**			**Value-Weighted Portfolios**		
	(2)	(3)	(4)	(2)	(3)	(4)		(2)	(3)	(4)	(2)	(3)	(4)
Mkt-rf	0/11	0/11	0/11	0/11	0/11	1/11	Mkt-rf	0/11	2/11	0/11	1/11	1/11	1/11
SMB	4/11	2/11	4/11	3/11	3/11	2/11	SMB	0/11	0/11	0/11	0/11	0/11	0/11
HML	1/11	0/11	4/11	0/11	0/11	0/11	HML	1/11	3/11	1/11	1/11	1/11	3/11
WML		3/11			4/11		WML		2/11			3/11	
RMW			2/11			1/11	RMW			1/11			0/11
CMA			2/11			3/11	CMA			0/11			2/11

This table summarizes the findings of the 11 industry analyses, using the quartile portfolios. It presents the fraction of significant risk factors for each model. Factors of all three significance levels, 1%, 5% and 10%, are counted. For example, looking at the period 1 equally-weighted portfolios, we find that for model 2, the HML factor is significant for one industry.

From the universal industry analysis, we recall finding more significance in the value-weighted portfolios than in the equally-weighted portfolios. When we split the data set in two periods, this pattern disappears, and the significance is more equally distributed between the two portfolio types. In other words, looking at the two periods, we do not find the same evidence for a difference in risk exposure between large and small firms, as we did in the universal industry analysis.

3.2.2 Summary Periodic Analysis

To sum up, in this section we have investigated whether splitting the sample into two time periods gives similar results as the universal analysis, or whether there are differences between the two periods. As in the universal analysis, the majority of the alphas are still negative when we split the sample. However, we discovered that there is more significance in the period 2009-2013, than in the period 2014-2018. Additionally, significance only appeared in the equally-weighted portfolios in period 2. This might be a result of increased focus and action aimed at ESG issues. In the first period, low ESG firms outperform high, by approximately 0.4%-1.6%. In period 2, small companies with low ESG scores outperform small companies with high ESG scores, by 0.5%-1%. Regarding the industries, we still observe insignificance for Energy, Real Estate and Technology. This reinforces the

assumption of a neutral relationship in these industries. Finally, the findings in Health Care suggested that the outperformance of high ESG firms by low only started in the second period. We now continue to evaluate the performance of our strategy in the United States and Europe.

3.3 Geographic Analysis

In this section, we perform the analysis on two different geographic areas, the United States and Europe. The reason behind this is that companies from the same countries often are more comparable because they face the same market conditions and governmental requirements. Not including other regions may be considered a limitation of the thesis. The lack of research on other parts of the world would also have made it interesting to study for example Asia and Africa. However, since prior research mainly has focused on these two regions separately, we find it interesting to examine the two areas together. Moreover, the US and Europe are together with Japan, the three largest regions in the world, based on the value of their sustainable investing assets (Global Sustainable Investment Alliance, 2018). Additionally, the two regions share many of the same characteristics and are often compared on different matters. Therefore, we find it interesting to explore how they compare on the matter of ESG investing.

From the initial screened sample of 7 711 companies, we extract companies from the US into one data set and companies from Europe into another data set. The American data set is larger than the European, with 2 426 against 1 075 companies. The analysis follows the same structure as sections 3.1 and 3.2. The aim is to investigate whether the results from the geographic analysis can explain the findings from the universal analysis, and whether there are differences in the results from the two regions. The main focus is still the alpha, which measures the abnormal return. Consequently, this is the only number we report in table 3.8. When it comes to portfolio size, we use 25% and 40% thresholds for the overall analysis. As in the periodic analysis, we reduce the number of models to three. The three excluded models, models 1, 5 and 6, showed similar results to models 2 and 4. To increase the readability of the tables and make the analysis easier to follow, we therefore decide to exclude these models. The results from the excluded models can be provided upon request.

Table 3.8: Geographic Analysis

Panel A: The United States							Panel B: Europe						
I: Portfolio size 40%							I: Portfolio size 40%						
	Equally-Weighted Portfolios			Value-Weighted Portfolios				Equally-Weighted Portfolios			Value-Weighted Portfolios		
	(2)	(3)	(4)	(2)	(3)	(4)		(2)	(3)	(4)	(2)	(3)	(4)
Mkt-rf	0.005	0.003	0.016	-0.027	-0.025	-0.001	Mkt-rf	-0.063***	-0.052**	-0.062**	-0.110**	-0.089	-0.147**
SMB	0.080	0.079	0.090	0.064	0.065	0.079	SMB	-0.063	-0.054	-0.062	0.009	0.026	0.019
HML	0.029	0.021	-0.022	0.027	0.036	-0.085	HML	0.086	0.120*	0.073	0.132	0.201	0.493**
WML		-0.011			0.012		WML		0.056			0.107	
RMW			0.038			0.049	RMW			-0.020			0.463
CMA			0.097			0.233*	CMA			0.013			-0.270
$\boldsymbol{\alpha}$	-0.307***	-0.307***	-0.339***	-0.464***	-0.464***	-0.526***	α	-0.084	-0.119	-0.077	-0.275	-0.345	-0.380
N	132	132	132	132	132	132	N	1321	132	132	132	132	132
R2	0.027	0.028	0.038	0.007	0.008	0.003	R2	0.059	0.076	0.059	0.034	0.048	0.067
Adj. R2	0.004	-0.003	-0.0002	-0.016	-0.023	-0.006	Adj. R2	0.037	0.047	0.022	0.011	0.018	0.030
II: Portfolio size 25%							II: Portfolio size 25%						
	Equally-Weighted Portfolios			Value-Weighted Portfolios				Equally-Weighted Portfolios			Value-Weighted Portfolios		
	(2)	(3)	(4)	(2)	(3)	(4)		(2)	(3)	(4)	(2)	(3)	(4)
Mkt-rf	0.009	0.006	0.032	-0.026	-0.027	0.005	Mkt-rf	-0.080***	-0.067**	-0.082**	-0.006	0.009	0.020
SMB	0.125*	0.123	0.136*	0.098	0.097	0.118	SMB	-0.026	-0.016	-0.032	-0.018	-0.005	0.093
HML	0.038	0.024	-0.079	-0.024	-0.026	-0.167	HML	0.110	0.152*	0.092	0.036	0.086	0.327
WML		-0.021			-0.003		WML		0.065			0.078	
RMW			0.029			0.058	RMW			-0.044			0.624**
CMA			0.219*			0.291*	CMA			-0.016			0.180
α	-0.336**	-0.337**	-0.392**	-0.417**	-0.417**	-0.494***	α	-0.055	-0.097	-0.036	-0.017	-0.067	-0.269
N	132	132	132	132	132	132	N	132	132	132	132	132	132
R2	0.033	0.035	0.057	0.011	0.011	0.042	R2	0.053	0.068	0.053	0.001	0.009	0.039
Adj. R2	0.010	0.004	0.019	-0.013	-0.020	0.004	Adj. R2	0.030	0.028	0.015	-0.023	-0.022	0.001

Note: *p<0.1; **p<0.05; ***p<0.01
This table presents the results from the geographic analysis. Panel A shows the results from the United States, while Panel B shows the results from Europe. For both regions we use 40% and 25% ESG thresholds. The models used in this analysis is the Fama French three factor model (2), the Fama French three factor model with momentum (3) and the Fama French five factor model (4).

Since we extract two regions from the global sample, it is worth mentioning that the portfolios are somewhat less diversified than in the universal and periodic analyses. Table 3.8 presents the results from the overall analysis for the US and Europe. First, we notice that all the alphas are negative, which is consistent with the previous findings. Secondly, the alphas in Europe are on average much closer to zero than in the US. The most interesting observation is however that only the US exhibits significant alphas. All alphas in the US panel, regardless of portfolio construction or model, are significant at the 1% or 5% level. These results suggest that investors following the strategy of buying high and selling low ESG companies in the US can expect to lose money. Low ESG companies outperform high by 0.3% to 0.5% monthly. Examining the European panel we find no significant alphas. This indicates that there is no significant difference in returns between high and low ESG companies in Europe. An interpretation of these results could be that the US is lagging behind Europe when it comes to incorporating actions towards ESG issues. This might be explained by the less significant investor focus on ESG in the US

compared to Europe. According to a survey by RBC Global Asset Management, the focus on ESG issues among American institutional investors is less than among European institutional investors. Furthermore, the survey suggests that American investors to a lesser extent than European investors are forced to incorporate ESG in their investment approach (RBC Global Asset Management, 2019). In Europe, the European Union and individual countries are moving forward with specific directives, with a focus on ESG incorporation into investments. In the US on the other hand, guidance for private-sector retirement plans stresses that fiduciaries must not put ESG goals ahead of financial goals. This might also affect the incentives businesses have to incorporate actions towards ESG issues, and can therefore contribute to explain our findings (Blackrock Investment Institute, 2019).

Turning to the risk factors, we again observe little significance. There is however more significance for Europe, indicating that there is a bigger difference in risk exposure between high and low ESG performing companies in Europe, than in the US. This finding may be a result of Europe being a region of different countries, which creates larger differences between European companies than American.

3.3.1 Geographic Industry Analysis

Next, we turn to the industry level analysis for the US and Europe. In this analysis, we only use portfolios constructed on the top and bottom 40% ESG performing companies. The reason for this is that the number of companies in the US and Europe naturally is lower than for the entire universe. Therefore, when we divide the samples further into industries, some industries become too small to use quartile portfolios. The yearly average number of companies for each industry is presented in table 3.9.

Table 3.9: Yearly average number of companies per industry in the United States and Europe

	The United States	**Europe**
Basic Materials	52	63
Consumer Discretionary	206	139
Consumer Staples	55	52
Energy	73	51
Financials	199	160
Health Care	134	51
Industrials	195	162
Real Estate	86	48
Technology	121	34
Telecommunications	31	37
Utilities	46	35
Total	1 198	832

This table provides the yearly average number of companies in each industry, for the United States and Europe.

Table 3.10 provides the industry alphas from the analysis of the United States and Europe. As expected from the findings in the overall geographic analysis above, we find more significant alphas in the US than in Europe. Further, there are more positive, though not significant, alphas in Europe. This is also consistent with the overall analysis, where the European alphas were close to zero, and contributes to explain the overall insignificance in Europe.

Table 3.10: Industry specific alphas for the United States and Europe

Panel A: The United States							**Panel B: Europe**						
	Equally-Weighted Portfolios			**Value-Weighted Portfolios**				**Equally-Weighted Portfolios**			**Value-Weighted Portfolios**		
	(2)	(3)	(4)	(2)	(3)	(4)		(2)	(3)	(4)	(2)	(3)	(4)
Basic Mat.	-0.089	-0.096	-0.085	-0.428	-0.431	-0.511	Basic Mat.	0.053	0.039	0.124	0.083	0.029	0.225
Con. D.	-0.377*	-0.379*	-0.429**	-0.568*	-0.568*	-0.674**	Con. D.	0.035	-0.141	0.036	-0.457	-0.561*	-0.434
Con. S.	-0.630**	-0.627**	-0.825***	-0.589*	-0.588*	-0.725**	Con. S.	-0.231	-0.183	-0.312	-0.613*	-0.725**	-0.853**
Energy	-0.581*	-0.575*	-0.611*	-0.446	-0.438	-0.500	Energy	0.568*	0.529	0.598	-0.207	-0.343	-0.067
Financials	0.011	0.006	0.059	0.043	0.035	0.125	Financials	-0.288	-0.236	-0.262	-0.276	-0.169	-0.249
Health Care	-0.694***	.9,693***	-0.826***	-0.975***	-0.977***	-1.168***	Health Care	0.330	0.304	0.176	0.477	0.452	0.087
Industrials	-0.107	-0.107	-0.102	-0.304	-0.304	-0.318	Industrials	-0.181	-0.199	-0.270	-0.293	-0.181	-0.259
Real Estate	-0.223	-0.223	-0.176	-0.199	-0.199	-0.136	Real Estate	0.359	0.411	0.478	-0.007	0.016	0.004
Technology	-0.169	-0.168	-0.217	-0.510	-0.509	-0.684*	Technology	0.364	0.344	0.424	-0.165	-0.345	-0.290
Telecom.	-0.551	-0.551	-0.597	-0.930**	-0.930**	-1.042**	Telecom.	-0.634**	-0.709**	-0.557*	-0.526	-0.549	-0.614
Utilities	-0.076	-0.077	-0.158	-0.144	-0.145	-0.176	Utilities	0.300	0.240	0.296	0.187	0.005	-0.338

Note: *p<0.1; **p<0.05; ***p<0.01
This table provides the alphas from the geographic industry level analysis. Panel A shows the alphas from the United States, while Panel B shows the alphas from Europe. For both periods we analyze equally- and value-weighted portfolios. All portfolios are constructed using a 40% ESG threshold. The applied models are the Fama French three factor model (2), the Fama French three factor model with momentum (3) and the Fama French five factor model (4).

From the universal industry analysis, we recall concluding that four industries showed

no difference in performance between the high and low portfolios. For Financials, Real Estate and Technology this does not change when we analyze the US and Europe. This supports the assumption that there is a neutral relationship between ESG and return in these industries. Regarding the fourth industry, Energy, the picture has changed slightly. According to the US analysis on equally-weighted portfolios, low ESG Energy stocks generate higher returns than high ESG Energy stocks. However, we notice that the level of significance is low, hence there is still uncertainty related to this assumption.

Basic Materials, Industrials and Utilities all showed several significant alphas in the universal analysis, though the majority only at the 10% significance level. Analyzing the US and Europe, we find no significance. This would imply that there is no significant difference in performance between high and low ESG performing companies in these industries, in the US or Europe. Moreover, it means that the significance for Basic Materials, Industrials and Utilities in the universal industry analysis is likely to be driven by other parts of the world.

Regarding Consumer Discretionary, we observe significance for all models in the US analysis. This indicates that the US contributes to the negative alphas in the universal industry analysis. The US analysis also exhibits significant alphas for Consumer Staples, the majority at the 1% and 5% significance levels. This is interesting because the findings from the universal industry analysis demonstrate very weak significance. Consequently, it seems likely that the American Consumer Staples industry differs from the average global Consumer Staples industry. According to our findings, American low ESG companies generate on average 0.7% higher monthly returns than American high ESG companies. With regard to Europe, we also find significant negative alphas for Consumer Staples, but not to the same extent as for the US. Nevertheless, the results indicate that it is likely that low portfolios outperform high in Europe as well.

Telecommunications display significant alphas for all the equally-weighted portfolios in the universal industry analysis. From the European analysis, we observe the same results. For the US, the significance appears in the value-weighted portfolios and not the equally-weighted portfolios. This could imply that for the US there is a difference in performance between high and low ESG companies among big firms, whereas in Europe the difference exists for smaller firms. Based on these findings it seems like the US and Europe are both

somewhat responsible for the negative alphas for Telecommunications in the universal analysis.

Finally, the industry with the most significance in the universal industry analysis was Health Care. It is therefore interesting to observe the difference in results between the US and Europe. While there is no significant difference in return between high and low ESG Health Care companies in Europe, all alphas for the US sample are significant at the 1% level. According to our results, monthly returns of low Health Care companies are on average between 0.7% and 1.2% higher than returns of high Health Care companies in the US. This might be an indication of the Health Care industry in the US lagging behind the Health Care industry in Europe when it comes to ESG issues.

The industry level risk factors are summarized in table 3.11. Consistent with the overall analysis, we observe slightly more significant risk factors for Europe than for the US. The significant factors are overall relatively evenly distributed between the equally- and value-weighted portfolios. An interesting observation is that the momentum factor is significant for more than half of the industries in Europe for the value-weighted portfolios. For the equally-weighted portfolios, it is only significant for one industry. This indicates that the difference in exposure to momentum between high and low ESG portfolios in Europe is larger for big companies than for small. Hence there is also more risk premium in the momentum factor for large European companies than for small.

Table 3.11: Fraction of significant risk factors in the geographic industry analysis

Panel A: USA							**Panel B: Europe**						
	Equally-Weighted Portfolios			**Value-Weighted Portfolios**				**Equally-Weighted Portfolios**			**Value-Weighted Portfolios**		
	(2)	(3)	(4)	(2)	(3)	(4)		(2)	(3)	(4)	(2)	(3)	(4)
Mkt-rf	3/11	2/11	1/11	1/11	2/11	1/11	Mkt-rf	3/11	2/11	1/11	3/11	1/11	2/11
SMB	0/11	0/11	1/11	0/11	0/11	1/11	SMB	1/11	1/11	1/11	1/11	1/11	2/11
HML	2/11	1/11	5/11	1/11	1/11	2/11	HML	2/11	3/11	2/11	3/11	1/11	4/11
WML		3/11			2/11		WML		1/11			6/11	
RMW			0/11			3/11	RMW			0/11			2/11
CMA			3/11			3/11	CMA			2/11			2/11

This table summarizes the findings of the 11 industry analyses, for the United States and Europe, using the 40% portfolios. The table shows the fraction of significant risk factors for each model. Factors of all three significance levels, 1%, 5% and 10%, are counted. For example, looking at the United States equally-weighted portfolios, we find that for model (2), the HML factor is significant twice.

3.3.2 Summary Geographic Analysis

In summary, the geographic analysis has shown that in general, low ESG performing companies outperform high in the US, by 0.3%-0.5%. In Europe, the findings imply that the relationship is neutral. This could imply that the US is lagging behind Europe in incorporating actions towards ESG. Furthermore, the industry results for both the US and Europe are fairly consistent with the universal industry analysis for the majority of the industries. Consumer Discretionary, Consumer Staples and Health Care are particularly significant for the US, suggesting that an investor applying the strategy in these industries is likely to destroy value. For Financials, Real Estate and Technology, we find insignificance in both the US and Europe, which is in line with the neutral relationship we observed in the universal industry analysis.

4 Conclusion

In this thesis, we aimed to investigate whether investing in sustainable and responsible companies generates abnormal returns. We did this by testing a strategy of buying a portfolio of companies with high ESG scores and selling a portfolio of companies with low ESG scores. The basis of the analysis was a global data set of 11 years, from 2008-2018. From this data set, we performed three different analyses, evaluating the performance of the investment strategy on an overall and industry level. Our findings suggest that the relationship between ESG and financial performance clearly is not positive, and more negative than neutral. The negative relationship is especially evident globally for smaller companies, globally in the period 2009-2013, and in the United States. Further, Health Care stood out as exhibiting the most consistently negative relationship among the 11 industries. Real Estate and Technology on the other hand, demonstrated a neutral relationship. Our findings are relevant for investors seeking to invest sustainably, as they outline when and where such a strategy will destroy or not destroy value. For an investor who exclusively seeks positive returns, the profitable strategy seems to be betting against ESG. In the following, we will summarize our findings.

In the first analysis, we applied the strategy to the entire global data set. All alphas suggested a negative relationship, but significance only appeared for the equally-weighted portfolios. From this, we inferred that small companies with low ESG scores outperform small companies with high ESG score, by 0.4%-0.6% monthly. For larger companies, the findings indicate that there is no significant difference between the high and low portfolios. These results are interesting because they contribute to the global perspective on ESG investing, and because they provide valuable information for an investor seeking to invest in sustainable companies.

In the second analysis, we divided the data set into two time periods, with the purpose of investigating differences between the periods. This analysis also inferred that the low portfolios outperform the high. However, the first period, 2009-2013, exhibited more significance than the second period, 2014-2018. In the first period the low portfolios generated between 0.4% and 1.6% higher monthly returns than the high portfolios. We interpreted these findings as an indication of ESG investing becoming more profitable over

time, which would be positive news for conscious investors. This development might be a result of the increase in attention towards ESG issues we have experienced in recent years.

In the third analysis, we extracted companies from the United States and Europe into two different data sets. This analysis suggested that low ESG companies outperform high in the US, but not in Europe. American low ESG performing companies outperform high by 0.3%-0.5% monthly. In Europe, the results indicated that the relationship between ESG and financial performance is neutral. These findings are in line with the perception of the US lagging behind Europe in ESG incorporation. These findings are also valuable for an investor seeking to invest sustainably, or an investor seeking shareholder value exclusively. Further, the findings raise awareness of the fact there can be geographical differences in the performance of our strategy.

Furthermore, we evaluated the performance of the strategy on industry level globally, for the two sub-periods and for the US and Europe. Overall, the industries inferred the same as the general analyses, that low ESG firms outperform high. However, the industries demonstrated different results, and the results changed between the three analyses. Despite this, some industries displayed consistent results across both time periods and regions. Real Estate and Technology demonstrate a clearly neutral relationship. In Health Care on the other hand, low ESG performing firms seem to outperform high, by 0.6%-1.7%. For the remaining industries, the relationship moves between being negative and neutral, making it challenging to draw a general conclusion. However, this only emphasizes the importance of industry level information for investors seeking to invest in high ESG performing companies.

Moreover, the fact that we observe differences between industries, regions and time periods, might contribute to explain the inconsistency in existing literature. Sustainability and responsible investments are topics with many debaters who express many different opinions. It is therefore perhaps not a surprise that the relationship between ESG and financial performance is not the same today as it was 10 years ago. For the same reason, it is not given that the relationship will remain the same for the next 10 years. Governmental regulations and requirements on both businesses and investors are also factors that impact the relationship. As is the preferences of customers. Both these factors vary between countries and regions, which was visible in our findings. In summary, our thesis exemplifies

the importance of both the time period, geography and industry composition of the sample under investigation.

In terms of future research questions, one could consider using other measures of financial performance than the return. Secondly, it could be interesting to study the ESG score broken down to the three pillars, to see if one pillar is more important in explaining financial performance than the others. Further, since the US and Europe are dominating the existing literature, it would also be interesting to study the relationship in other parts of the world. Finally, there are different providers of ESG ratings, and an analysis of the degree of similarity between the agencies is also a relevant research question.

References

Baired, P., Geylani, P., and Roberts, J. (2012). Corporate Social and Financial Performance Re-Examined: Industry Effects in a Linear Mixed Model Analysis. *Journal of Business Ethics*, 109(3):367–388.

Blackrock Investment Institute (2019). Sustainability: The future of investing. URL:https://www.blackrock.com/us/individual/literature/whitepaper/bii-sustainability-future-investing-jan-2019.pdf, Accessed on 2019-12-16.

Chang, C. E., Nelson, W. A., and Doug Witte, H. (2012). Do green mutual funds perform well? *Management Research Review*, 35(8):693–708.

Climate Disclosure Standards Board (2019). Insights from the reporting exchange: Esg reporting trends. URL:https://www.cdsb.net/sites/default/files/cdsb_report_1_esg.pdf, Accessed on 2019-11-20.

Douglas, E., Van Holt, T., and Whelan, T. (2017). Responsible investing: Guide to ESG data providers and relevant trends. *Journal of Environmental Investing*, 8(1):91–114.

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.

French, F. R. (2019). Current research return. URL:https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html,.

French, K. R. (2019). Description of Fama/French 5 Factors for Developed Markets. URL:http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/f-f_5developed.html, Accessed on 2019-11-12.

Friede, G., Busch, T., and Bassen, A. (2015). ESG and financial performance: Aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment*, 5:210–233.

Galbreath, J. (2013). ESG in focus: The Australian evidence. *Journal of business ethics*, 118(3):529–541.

Global Sustainable Investment Alliance (2018). Global sustainable investment review 2018. URL:http://www.gsi-alliance.org/wp-content/uploads/2019/06/GSIR_Review2018F.pdf, Accessed on 2019-12-05.

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.

Hamilton, S., Jo, H., and Statman, M. (1993). Doing well while doing good? The investment performance of socially responsible mutual funds. *Financial Analysts Journal*, 49(6):62–66.

Kell, G. (2018). The Remarkable Rise of ESG. *Forbes*. URL:https://www.forbes.com/sites/georgkell/2018/07/11/the-remarkable-rise-of-esg/#54894e141695, Accessed on 2019-12-17.

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

Lean, H., Ang, W. R., and Smyth, R. (2015). Performance and performance persistence of socially responsible investment funds in Europe and North America. *The North American Journal of Economics and Finance*, 34:254–266.

MSCI (2019a). Esg 101: What is esg? URL:https://www.msci.com/what-is-esg, Accessed on 2019-12-17.

MSCI (2019b). The GLobal Industry Classification Standard (GICS). URL:https://www.msci.com/gics, Accessed on 2019-11-12.

Orlitzky, M., Schmidt, F. L., and Rynes, S. L. (2003). Corporate social and financial performance: A meta-analysis. *Organization studies*, 24(3):403–441.

Pastor, L. (2018). Liquidity Factors of Pastor and Stambaug (JPE 2003) updated through DEC 2018. URL:https://faculty.chicagobooth.edu/lubos.pastor/research/liq_data_1962_2018.txt?fbclid=IwAR1Pfyz9EGC60Ep9pm3SQZ0ublBbwtumjYWrsMehNZRzPYCCdNmDxUgadvw, Accessed on 2019-11-12.

Principles for Responsible Investments (2018). Principles for responsible investments. URL:https://www.unpri.org/download?ac=6303, Accessed on 2019-12-16.

RBC Global Asset Management (2019). An Evolving Landscape - 2019 Responsible Investment Survey Executive Summary. URL:https://www.rbcgam.com/documents/en/other/esg-executive-summary.pdf, Accessed on 2019-11-25.

Refinitiv (2019). Environmental, Social and Governance (ESG) Scores from Refinitiv. URL:https://www.refinitiv.com/content/dam/marketing/en_us/documents/methodology/esg-scores-methodology.pdf, Accessed on 2019-11-10.

Refinitiv (n.d.). Datastream - The World's Most Comprehensive Financial Historical Database. URL:https://www.refinitiv.com/content/dam/marketing/en_us/documents/fact-sheets/datastream-economic-data-macro-research-fact-sheet.pdf, Accessed on 2019-11-10.

Renneboog, L., Ter Horst, J., and Zhang, C. (2008). Socially responsible investments: Institutional aspects, performance, and investor behavior. *Journal of banking & finance*, 32(9):1723–1742.

Revelli, C. and Viviani, J. (2015). Financial performance of socially responsible investing (SRI): what have we learned? A meta-analysis. *Business Ethics: A European Review*, 24(2):158–185.

The Global Compact (2004). Who cares wins - connecting financial markets to a changing world. URL:https://www.unglobalcompact.org/docs/issues_doc/Financial_markets/who_cares_who_wins.pdf, Accessed on 2019-12-17.

Thomson Reuters (2017). Thomson Reuters ESG Scores. URL:https://www.esade.edu/itemsweb/biblioteca/bbdd/inbbdd/archivos/Thomson_Reuters_ESG_Scores.pdf, Accessed on 2019-11-13.

UNEP Finance Initiative (2019). Changing course. URL:https://www.unepfi.org/wordpress/wp-content/uploads/2019/05/TCFD-Changing-Course-Oct-19.pdf, Accessed on 2019-12-05.

Utgård, J. (2017). Er samfunnsansvar lønnsomt? *Magma*, 7.

Verheyden, T., Eccles, R. G., and 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.

Waddock, S. A. and Graves, S. B. (1997). The corporate social performance–financial performance link. *Strategic management journal*, 18(4):303–319.

Appendix

Table A0.1: Industry Classification

Industry	**Sub-industries**
Basic Materials	Chemicals, Metals & Mining
	Construction Materials
	Paper & Forest Products
	Containers & Packaging
Consumer Discretionary	Retailing
	Consumer Services
	Durables & Apperal
	Automobiles & Components
Consumer Staples	Food & Staples Retailing
	Food, Tobacco & Beverage
	Household & Personal Products
Energy	Oil, Gas & Consumable Fuels
	Energy Equipment & Services
Financials	Banks
	Insurance
	Diversified Financials
Health Care	Biotechnology
	Pharmaceuticals & Life Sciences
	Health Care Equipment & Services
Industrials	Transportation
	Capital Goods
	Commercials & Professional Services
Real Estate	Real Estate Management & Development
	Equity Real Estate Investment Trusts
Technology	Software & Services
	Technology Hardware & Equipment
	Semiconductors & Equipment
Telecommunications	Telecommunication Services
	Media & Entertainment
Utilities	Gas-, Water-, Electric-, and Mulit-Utilities
	Independent, Power and Renewable Electricity Producers

This table provides the sub-industries within each of the 11 GICS-industries.

Table A0.2: Universal analysis 40% portfolios

	Equally-Weighted Portfolios						**Value-Weighted Portfolios**					
	(1)	(2)	(3)	(4)	(5)	(6)	(1)	(2)	(3)	(4)	(5)	(6)
Mkt-rf	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	0.03	0.03	0.10	-0.03	-0.02	-0.02
SMB		-0.05	-0.05	-0.06	-0.06	-0.05		-0.10	-0.06	-0.33	-0.37	-0.38
HML		0.01	0.00	-0.02	-0.03	-0.03		0.06	0.27	-0.26	0.15	0.15
WML			-0.01		-0.01	-0.01			0.29***		0.35***	0.35***
RMW				-0.050	-0.05	-0.05				-0.98**	-1.09***	-1.09***
CMA				0.03	0.04	0.03				0.00	-0.42	-0.40
LIQ						-0.01						-0.02
α	-0.33***	-0.33***	-0.33***	-0.32**	-0.32**	-0.32**	-0.43	-0.43	-0.51	-0.10	-0.08	-0.08
N	132	132	132	132	132	132	132	132	132	132	132	132
R2	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.06	0.05	0.12	0.12
Adj. R2	-0.00	-0.02	-0.02	-0.03	-0.04	-0.05	-0.01	-0.02	0.03	0.02	0.08	0.07

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

This table present the results from the universal analysis, for the 40% top and bottom ESG performing companies. This table presents the results from the universal analysis. The table include equally- and value-weighted portfolios. The dependent variable is the excess return of the long-short zero-investment portfolio. The independent variables are interpreted as follows: Mkt-rf is the market risk premium, also referred to as the systematic risk. The SMB factor shows the portfolio 's exposure to small market cap stocks, while the HML factor is the exposure to high book-to-market stocks. The WML is referred to as the momentum factor, as it captures exposures to winners versus looser. RMW captures the exposure to companies with robust profitability, and CMA shows the portfolio's exposure to companies with a conservative investing strategy. The alpha is the intercept, and represents the abnormal return achieved by applying this strategy. The model is estimated using monthly data from 2008-2018.