Norwegian School of Economics
Bergen, fall 2018

The Effect of Corporate Sustainability on Stock Performance

An empirical comparison between European stocks with good and bad ESG performance

Simon Sveen Togba & Espen Ketilsønn Lyng

Supervisor: Carsten Gero Bienz

Master thesis in Financial Economics

NORWEGIAN SCHOOL OF ECONOMICS

Abstract

This thesis investigates the effect of corporate sustainability on financial performance in Europe during the period 2005-2017. We examine whether companies with good ESG performance perform better in the stock market than companies with bad ESG performance, based on Thomson Reuters ESG Scores. We are computing the alphas of a long-short zero investment strategy, which is long a portfolio comprised of companies with high ESG scores and short a portfolio comprised of companies with low ESG scores. By applying Fama-French three-factor, four-factor (Carhart) and five-factor model with and without momentum to account for potential differences in risk exposure between the portfolios, we find that the latter significantly outperforms the former. The differences however, disappear as we account for ESG controversies, which are the company´s involvement in media covered incidents related to ESG. Due to different results associated with the different ESG measures, a clear conclusion can hardly be made. What is certain however, is that we do not see a positive relationship between ESG and stock performance in Europe, using Thomson Reuters ESG Scores.

Preface

With this thesis, we complete our Master of Science in Finance at the Norwegian School of Economics, Department of Finance.

The increased attention on the topic of corporate sustainability has awakened our curiosity. We build on the existing research area and conduct an extensive analysis of the stock effects of corporate sustainability in Europe, to hopefully fill a gap in the literature.

Working with this thesis has been time consuming and challenging, yet highly rewarding. Throughout the work we have focused on learning and gaining knowledge on the topic of sustainable investments, and we are proud of what we have managed to learn and understand within the field.

We would like to express our gratitude to our supervisor, Carsten G. Bienz, for his generous support throughout the process. It has been crucial and given us the motivation necessary to complete this thesis. In addition, we would like to thank Torgeir Stensaker, Head of Fixed Income Norway at Nordea Investment Management AB, for inspiring us and introducing us to the concept of ESG.

Bergen, December 2018

Espen Ketilsønn Lyng

Simon Sveen Togba

Contents

1 Introduction

In recent decades, numerous investors have embraced the concept of socially responsible investing (SRI). The term refers to investment strategies that attempt to combine social and environmental benefits with financial return (McIntosh and Brzeszczyński, 2014). Globally, there were $22.89 trillion assets managed under responsible investment strategies at the beginning of 2016, compared to $18.28 trillion in 2014 (Global Sustainable Investment Alliance, 2016). Despite the increased popularity however, there is a continuous debate over whether adding social, environmental and ethical dimensions to the stock selection process add value or not. In other words, whether there exist a positive relationship between corporate social responsibility (CSR) and corporate financial performance (CFP). Friedman (1982) argues that the primary objective of managements is to maximize shareholder wealth, and that a company cannot use financial resources to improve corporate social performance (CSP) without destroying shareholder value. On the other hand, Cheng et al. (2014) claim that CSR can strengthen stakeholder relationships, and that firms with good CSP will have lower capital constraints through better access to bank loans, making it easier to undertake strategic investments. Others believe that the subject is complex and that there exist so many variables between the two, leaving no reason to expect a relationship (e.g. Ullmann, 1985).

Due to the debate, several studies have measured the link between CSR and CFP throughout the years. The majority have examined the performance of SRI funds (Lima, 2017), but this may not be the most appropriate option. The financial returns of mutual funds may be biased because of non quantifiable aspects, such as stock picking and timing ability of the fund management, which is difficult to separate from the financial performance of sustainable stocks. In addition, different screening methods cannot be analyzed separately, because mutual funds often employ multiple screens together. The results have consequently usually shown no significant difference in performance between SRI funds and conventional funds (e.g. Bauer et al.

2005). On stock level however, the majority of previous research have shown a positive linkage. Pava and Krausz (1996) reviewed 21 empirical studies and found that nearly all of these discovered that firms perceived as having met social responsibility criteria either have outperformed or performed as well as other firms. Orlitzky et al. (2003) made similar findings when evaluating a large number of studies from the 1970s to late 1990s. Nevertheless, the overall research area is not univocal as many researchers have failed to identify a positive relationship between the variables (e.g. Peng and Yang, 2014).

While most empirical studies have focused on the US, we narrow the investment universe to comprise 401 European companies from the STOXX 600 Europe index. The index consists of small, mid and large cap companies and we found this to be the most representative and diversified index for Europe as a whole. The reason why we chose this market is that despite increased focus on sustainable investments, the literature on SRI in Europe is deficient, and we believe that the research area will benefit from research input from this region to better understand the relationship between CSR and CFP.

We measure CSR by using ESG scores. ESG refers to the three central factors in measuring the sustainability of a company; environmental, social, and governance, and the concept has rapidly grown to be the general category dominating the SRI landscape (The Forum for Sustainable and Responsible Investment, 2016). In 2017, 85% of European institutional asset owners, compared to 67% worldwide, used ESG principles as part of their investment approach, according to a survey carried out by RBC Global Asset Management (2017). Due to this popularity in Europe, we find ESG to be the natural CSR measurement approach. Unlike SRI ratings, which use mostly negative screens, such as not investing in sin companies[1] ESG integration is based upon the assumption that ESG factors, can have a material impact on the

[1]Companies involved in activities that are considered inconsistent with social norms, e.g. alcohol, tobacco, firearm and gambling companies.

value of securities and the long-term performance of companies.

The ESG scores we use are carried out by Thomson Reuters. Primarily because they have the most historical ESG data on European companies, being the first rating agency to provide raw ESG data (Polk and Llp, 2017). At the same time, they evaluate companies on more key issues than the other ESG data providers (Douglas et al., 2017), and by doing this, cover the multidimensional concept of CSR in a better way. In addition, Thomson Reuters are the only of the main ESG players that do not rely on company input in their assessment of ESG performance (Douglas et al., 2017). We believe this makes them more objective, and improves the reliability of their scores.

In this thesis, we examine whether companies with good ESG performance perform better in the stock market than companies with bad ESG performance. ESG performance cover ESG scores, which are calculated on the basis of company-reported data, and ESG combined (ESGC) scores, which overlays the ESG score with ESG controversies to magnify the impact of significant controversies. ESG controversies are corporate ESG news stories such as suspicious social behaviour and resource use scandals that place a firm under the media spotlight and, by extension, grab investors' attention (Cai et al. 2012). The effect of controversies are based on an assessment of each company´s involvement in media covered incidents related to ESG. The controversies are not controlled by the company since they are disclosed by the media, and by applying this score as well as the ESG score, we are able to assess the overall sustainability performance of companies.

We sort the companies into portfolios, using deciles and quartiles, hereby called small and large portfolios. The portfolios are further characterized as "good" or "bad" on the basis of their ESG score. We repeat the sorting process, but this time on the basis of ESGC score, making up a total of eight different portfolios. They rebalance every month for the whole time span of 13 years, so we can be certain that the portfolios consist of the best and worst ESG performers at all times. As a robustness test, we also rebalance the port-

folios on a yearly basis. To measure the difference in stock performance, we compute the alphas of a long-short zero investment strategy which is long the good portfolios and short the bad portfolios. We apply Fama-French three-factor, four-factor (Carhart) and five-factor with and without momentum to account for potential differences in risk exposure between the groups, and also to capture any cross sectional differences.

We track stock performance for 13 years and find that the bad portfolios outperform the good portfolios consistently throughout our models, when sorted on the basis of ESG score. For the small portfolios, the monthly difference is on average around 0.3-0.4% per month and for the large around 0.5% per month. Theoretically, this implies that abnormal returns can be made by investors applying a long-short strategy, being long companies with a low ESG score and short companies with high ESG score. However, the outperformance vanishes when the portfolios are sorted on the basis of ESGC score. These scores fluctuate more than the ESG scores, and the turnover of stocks coming in and out of the portfolios as they rebalance every month is higher. ESG controversies seem to hurt stocks more randomly, making it harder for investors to formulate any strategy on the basis of this score. Furthermore, our results show that the companies in the good portfolios, sorted by ESG scores, are larger than the companies in the bad portfolio.

We believe that our results form a complex image of SRI in Europe, and due to different results associated with the different ESG measures, a clear conclusion can hardly be made. What is certain however, is that we do not see a positive relationship between ESG and stock performance for the STOXX 600 Europe, using Thomson Reuters ESG Scores.

The remaining parts are structured as follows: Part 2 provides an overview of related literature, while part 3 contain our hypotheses for this thesis. Part 4 elaborates on the data and assumptions used, whereas part 5 describes the methodology. Part 6 reveals the results, before part 7 and 8 respectively presents the discussion and a final conclusion.

2 Literature review

In the following section we will discuss theoretical arguments and the empirical findings presented in the literature. We will also, on the basis of previous research, explain why we use ESG to measure CSR.

2.1 Theoretical background

The current theory regarding the link between SRI and CFP remains ambiguous. Arguments for positive, negative and neutral relationships have all been presented throughout the years. Those arguing for a negative relationship claim that socially responsible companies have a competitive disadvantage (e.g. Aupperle et al., 1985) since they are incurring unnecessary costs. The belief is that processes and practices that aim to achieve a high sustainability profile only provide additional costs or inefficient resource allocation. The argumentation is in line with the arguments of neoclassical economists like Friedman (1970) who believes that the cost of being socially responsible outweighs the benefit, and therefore contributes to reducing profits and shareholder wealth. He believes that a company cannot use financial resources to improve CSP without destroying shareholder value.

Moreover, slack resource theory postulates that it is the resources owned by a company that enables them to adapt to internal or external pressure for change (Fauzi and Idris, 2009). The theory claims that strong financial performance provides companies with the opportunity to act socially responsible. The reasoning being that if a company improves their financial performance, there will be resources available, which allows the company to engage in socially responsible activities, like community and employee relations (Waddock and Graves, 1997). Similarly, Roberts (1992) and Ullmann (1985) state that companies with bad financial performance seek investment opportunities with shorter time horizon and more immediate results rather than socially responsible investments.

Another theory often used to explain the relationship between CSR and CFP is the good management theory (Miles and Covin, 2000). This theory argues that the reason there would exist a relationship between the two is because good social and environmental performance can improve relationships with key stakeholder groups (Cheng et al., 2014). Examples are good employee relations which is expected to improve productivity, lower capital constraints through better access to bank loans, and good community relations which may incentives local government to reduce regulation, hence costs, and therefore the financial performance. Accordingly, Miles and Covin (2000) claimed that social and environmental performance could be an alternative way to satisfy stakeholders and provide a competitive power. Another argument for a positive relationship stipulates that CSR can work as an indicator for superior management skills, and thus lead to lower explicit costs (Alexander and Buchholz, 1978).

Those that believe no relationship exist believe that there are so many variables existing between CSR and CFP, leaving no reason to expect a relationship (e.g., Ullman, 1985). Other arguments for a neutral relationship revolve around the problems associated with measuring CSR, and that these may disguise any connection between the two (Turker, 2009).

2.2 Empirical evidence

Ever since the subject of CSR became popularized in the 1970s, the possible linkage to CFP has been of increased interest of researchers. Many empirical studies have found a positive relationship, but researchers often claim that the results are inconclusive or contradictory (Aupperle et al., 1985; Griffin and Mahon, 1997). The previous studies can be divided into three levels of asset analysis: Funds, stocks and indices. In this section we review the empirical findings in the literature on SRI funds and SRI stocks, since they are the most relevant to our research.

2.2.1 SRI funds

The research on SRI funds is the most commonly researched SRI asset level (Lima, 2017) and usually involves comparing socially responsible funds with conventional funds. The initial evaluation of the performance of these funds was made by Hamilton et al. (1993). By using the CAPM (1964) and Jensen's alpha (1968), they compared the performance of 32 SRI funds with a benchmark made up of 170 conventional funds in the period of 1981-1990. They found no significant difference between the two groups and concluded that social responsibility factors did not have any effect on expected return on stocks. Mallin et al. (1995) who used a matched pair approach, matching UK SRI funds and conventional funds by size and age in the period of 1986-1993, neither found any statistically significant difference in performance, using the Sharpe (1966), Treynor (1965) and Jensens alpha measures. These findings are later confirmed by Statman (2000), applying the same method on US funds during the period 1990-1998.

The prior studies applying single-factor models met some critics for not being able to capture cross sectional differences (Fama and French, 1992, 1993, 1995, 1998). Therefore the recent literature to a larger extent applies multi-factor models like Fama-French (1993) and Carhart (1997). Bauer et al. (2005) apply Carhart four-factor model to analyze the performance of SRI funds in UK, US and Germany during the period 1990-2001. Using this approach they overcame benchmark problems that single factor models could not, but the results still showed insignificant difference between these and conventional funds. They later (2006, 2007) extended the research to include Australian and Canadian funds but still could not find any statistically significant results. Renneboog et al. (2008) expanded the study even further to include funds all over the world in the period 1991-2003. They found that most SRI funds performed worse than their conventional domestic benchmark, however, when adjusting for risk they found that the performance of the two groups was statistically no different, with only a few exceptions.

2.2.2 SRI stocks and portfolios

The findings on SRI stock performance have in general differed from the findings on SRI fund performance. The earliest research on SRI stocks and portfolios was performed by Vance (1975) who found a negative relationship between CSR and CFP tracking US companies over a period of three years. However, the results have in most cases been positive in the years following that. Orlitzky et al. (2003) reviewed a large number of studies looking at the performance of SRI stocks from the 1970s to late 1990s, and found that CSR is positively correlated with performance in the stock market. In addition to stock performance, they included accounting measures like return on assets (ROA) and return on equity (ROE), and came to the conclusion that managers should attend to be socially responsible, since the market later will reward them for it.

More recent studies applying multi-factor models generally show the same results. Kempf and Osthoff (2007) compared several screening strategies and found that large abnormal returns could be made by buying US stocks with a high score on KLD Research & Analytics and selling US stocks with a low score. During the period of 1992-2004 they found that investors could actually earn as much as 8.7% per year applying this strategy. Statman and Glushkov (2008) later performed an analysis using the Carhart-model on US companies in the period 1992-2007, and similarly found that companies with a high score from KLD Research & Analytics performed better in the stock market than conventional stocks with a low score. However, they also learnt that this was mostly offset by the advantage of not investing in sin stocks, which supports Hong and Kacperczyk's (2009) study, where they found that these companies get rejected by many investors.

New research done by Eccles et al. (2014) made findings in line with previous research when they studied a matched sample of 180 companies in the period 1993-2010. Out of the 180 companies, 90 were classified as high sustainability companies and 90 classified as low sustainability companies, based on a num-

ber of environmental and social policies adopted in the early 1990s, reflecting strategic choices and not the recent hype around sustainability issues (Eccles and Krzus, 2010). Using a Fama–French three-factor model augmented by the Carhart momentum factor they found that the high sustainability companies outperformed the low sustainability group in the stock market as well as in accounting measures.

To summarize, there exist little evidence of SRI funds over or underperforming relative to the market, and also little evidence of differences between SRI funds and conventional funds. However, on stock level, even though some studies have failed to identify a positive relationship (e.g., Peng and Yang, 2014), there seem to have been established a positive link between CSR and CFP, particularly through Orlitzky et al. (2003).

2.3 Measurement of CSR

There have been used numerous of different methods to measure CSR in the recent decades. The most commonly used are content analysis of corporate communication, questionnaire-based surveys, one dimensional measures and measurement via reputation indices and scales (Galant and Cadez, 2017). In this study we use the last measurement method, and more precisely ESG scores, as we believe that the other measurements are victim to some major weaknesses. Content analysis will to a large extend be affected by the researchers subjectivity, questionnaire-based surveys are exposed to response biases and one-dimensional methods is theoretically problematic since CSR evidently is a multidimensional issue (Galant and Cadez, 2017).

2.3.1 ESG

ESG is a modern investment term often used simultaneously as sustainability and refers to the three central factors in measuring the sustainability of a company, which are environmental, social and governance. ESG factors are used to enhance traditional financial analysis by identifying investment opportunities beyond regular technical valuations, and The Forum for Sus-

tainable and Responsible Investment (2016) reported that ESG incorporation has grown to be the general category dominating the SRI landscape. While less than 20 companies disclosed ESG data in the beginning of the 1990s, the number has increased to nearly 9000 (Geraghty and Vanderzeil, 2017). The belief that taking corporate environmental, social and governance risk factors into account, can improve financial returns, has also gained attention across capital markets all around the world. Several pension funds and insurers have started to award new business solely to asset managers with ESG capabilities (Kell, 2018). Additionally, RBC Global Asset Management (2017) found that 67% of institutional asset owners used ESG principles as part of their investment approach and decision making. In Europe, the amount was 85%, which is more than in any other region. Due to this popularity, especially in Europe, we believe that ESG is the most relevant measure of CSR, for our thesis. Besides, since ESG is a rather new phenomenon, there have been relatively few studies examining the effect ESG ratings might have on stock performance.

3 Hypothesis

This section introduces our main hypothesis as well as our supplementary research question. They are mostly based on previous empirical results, but also our own assessment.

3.1 Main hypothesis

Companies with good ESG performance perform better in the stock market than companies with bad ESG performance

We expect companies with good ESG performance to have significantly higher risk adjusted returns than companies with bad ESG performance. Previous research on SRI stock performance has usually presented similar findings (e.g., Orlitzky et al., 2003), especially in the US, and we expect the same results to be shown in Europe. Even though relatively few studies have measured CSR by using ESG performance, we assume that the findings will be similar.

3.2 Supplementary research question

What are the drivers of ESG and the possible differences in stock performance?

This question is much more complex than our main hypothesis, and is not a question that we necessarily aim to answer in this thesis, but rather use as a basis for discussion throughout. Based on the results from our analysis, we will discuss the question and form alternative explanations with the purpose of motivating for further research.

4 Data

This section describes the process of collecting the data and making it ready for the analysis. It provides descriptions of our data sources, how the data is generated, and how the ESG scores are calculated and carried out by Thomson Reuters. Furthermore, we explain the sample selection process and the variables used in the model.

4.1 Data sources

The data is collected from the Thomson Reuters Datastream and Kenneth R. French's data library. Datastream is a global financial and macroeconomic data platform covering equities, stock market indices, currencies, company fundamentals, fixed income securities and key economic indicators for 175 countries and 60 markets (Reuters, 2018a). From Datastream we retrieved industry membership, market value, bond yields, exchange rates, adjusted stock prices as well as the ESG performance. We use Kenneth R. French's data library to download the factors used in the Fama French asset pricing models.

4.2 Thomson Reuters ESG Scores

Thomson Reuters ESG Scores was launched in May 2018 and is an upgrade and replacement to the widely used ASSET4 database (Reuters, 2018c). The database contains ESG scores on over 7000 companies across the world, with time series data going back to 2002. Among the main ESG data providers, Thomson Reuters have the most historical ESG data on European companies, being the first to provide raw ESG data to investors (Polk and Llp, 2017). At the same time they evaluate companies on more key issues than any other ESG data provider (Douglas et al., 2017). In addition, they are the only one that does not rely on company input in their assessment of ESG performance (Douglas et al., 2017). We believe this improves their objectivity, as well as the reliability of their scores. Thomson Reuters ESG database is also substantially less costly compared to its closest substitutes, which makes the

data more available to private investors, as well as institutional investors. In order to assess overall ESG performance, they offer two overall ESG scores, Thomson Reuters ESG score and Thomson Reuters ESG Combined (ESGC) score.

4.2.1 ESG score

The ESG score is calculated using over 400 measures, which are based on considerations around industry relevance, compatibility and data availability (Reuters, 2018c). Of the 400 measures, the 178 most comparable and relevant are used in the final scoring process. They are further grouped into 10 main categories, which are; Resource use, emissions, innovation, management, shareholders, CSR strategy, workforce, human right, community and product responsibility. The score in each of the main categories, proportionately weighted to the count of measures in each category, is the basis for the three pillars, Environmental, Social and Governance. Detailed weights and counts are presented in table 1 below:

Table 1

Pillar	Category	Indicators in rating	Weights	Pillar Weights
Environmental	Resource Use	19	11 %	(11%+12%+11%)
	Emissions	22	12 %	
	Innovation	20	11 %	
Social	Workforce	29	16 %	(16%+4.5%+8%+7%)
	Human Rights	8	4.50 %	
	Community	14	8 %	
	Product Responsibility	12	7 %	
Governance	Management	34	19 %	(19%+7%+4.5%)
	Shareholders	12	7 %	
	CSR Strategy	8	4.50 %	
TOTAL		178	100 %	

This table shows how Thomson Reuters has calculated the ESG scores and how the main categories are weighted (Reuters, 2018c).

4.2.2 ESG score calculation methodology

The scores are calculated using a percentile rank scoring methodology (Reuters, 2018c). The calculation is based on assessments of the company's ESG performance relative to the others, and will depend on how many companies that is worse than the current one, how many that have the same value and how many that do not have a value at all. The scores range from 1-100 and is the equally weighted sum of all relevant industry indicators, excluding quantitative indicators that do not have publicly available data. The scores are calculated as follows:

$$\frac{\#\text{ of companies with a worse value} \frac{\#\text{ of companies with the same value including the current one}}{2}}{\#\text{ of companies}} \tag{1}$$

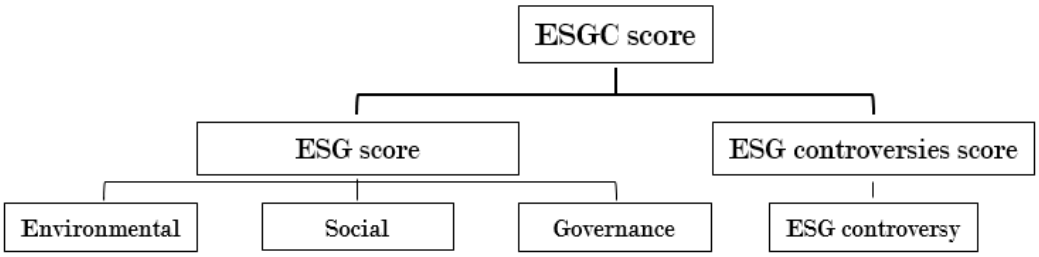
4.2.3 ESG Combined score (ESGC score)

The main objective of the ESGC score is to discount the ESG score based on negative media attention (Reuters, 2018c). It does so by incorporating the impact of significant, material ESG controversies. ESG controversies are corporate ESG news stories that place a firm under the media spotlight, e.g. events related to employee health, tax frauds, customer safety or the environment. The scores are dependent on each company´s number of involvements in incidents related to ESG, and to assess controversies, Thomson Reuters applies an individual controversies score. The controversies score is calculated using 23 ESG controversy measures[2] where recent controversies are reflected in the latest period. If a company is not involved in any controversies, the ESGC score will equal the ESG score, but when a company has been involved in a controversy, the ESGC score will equal the weighted average of the ESG Score and the ESG controversies score. If a scandal occurs during the year, the company involved is penalized in the ESGC score, and if the impact of the scandal is still seen in the following year, e.g. through

[2]The 23 ESG controversy measures used to calculate the controversies score are presented in the appendix table 14.

lawsuits or ongoing fines, it will further impact the score in this period. The following figure shows the relationship between the different scores:

Figure 1



This figure shows how the ESG and ESGC score are connected.

4.3 Sample selection

Our data sample consists of European stocks from the STOXX Europe 600 index. This index represents 600 large, mid and small capitalization companies across 17 countries in the Western Europe (STOXX, 2018), and we found this to be the most representative and diversified index of European stocks. The reason why we chose European companies is the lack of research on the performance of sustainable stocks in this market. Previous studies have mainly focused on US stocks, and we believe that the research area will benefit from research input from the European region.

4.3.1 Screening

First, we needed to make sure we had sufficient historical data for our time series analysis. We eliminated companies with less than 13 years of continuous ESG data, which left us with 409. Second, we eliminated the companies that did not have sufficient historical data on adjusted stock prices, which made us remove another eight. The 401 companies left in the data sample will hereby be referred to as "the index". The industry composition in the index remains similar to the complete STOXX 600 index and is shown in the following table:

Table 2

Sector composition of the index					
	Companies	Value weighted		Companies	Value weighted
Industry	% of index	% of index	Industry	% of index	% of index
Oil & Gas	4.24	8.22	Cons. Services	10.72	6.51
Basic Materials	7.48	6.02	Telecom	3.49	6.15
Industrials	21.20	11.07	Utilities	5.49	6.35
Cons. Goods	12.72	17.00	Financials	24.19	24.16
Health Care	6.98	11.34	Technology	3.49	3.18
Total				100.00	100.00

This table shows the sector composition in the index, both equally weighted and value weighted

4.3.2 Picking portfolios

We divide the 401 companies into portfolios sorted by each company's ESG performance, one grouping based on ESG score and one based on ESGC score. For the main analysis, the portfolios rebalance every month for the whole time span of 13 years, so that they consist of the best and the worst ESG performers at all time. We are performing the analysis on top and bottom quartiles as well as the top and bottom deciles, thus we need to make 1,248 different portfolios to perform this analysis. For robustness, we repeat the process with yearly rebalancing portfolios instead of monthly.

The behaviour of the monthly rebalancing portfolios based on ESG score differs a lot from the ones based on ESGC score. In terms of score fluctuation, the ESGC scores have a 36% higher average standard deviation than the ESG scores. This volatility causes the average monthly turnover of stocks to be 78% higher for the portfolios based on ESGC score than the ones based on ESG score. Also, over the total time period, the number of different stocks that have been a part of the ESGC portfolios, is 46% higher than for the ESG portfolios.[3]

[3]See table 19 in the appendix for details

4.4 The variables

For our dependent variable, we use monthly returns of a long-short zero investment[4] portfolio which is long the good portfolio and short the bad portfolio. The stock prices used in the calculation of returns are the closing prices adjusted for subsequent capital actions (Reuters, 2018b). The monthly stock performance is calculated as a m/m rate of change:

$$r_t = \frac{P_t}{P_{t-1}} - 1 \tag{2}$$

$$\begin{aligned} &Where: \\ &\quad r_t \;=\; Return\ at\ time\ t \\ &\quad P_t \;=\; Adjusted\ stock\ price\ at\ time\ t \end{aligned}$$

Furthermore, we use both the equally weighted return and the value weighted return of the long-short portfolios. The value weighted return is calculated as follows:

$$r_{pt} = \sum_{i=1}^{N}(w_{it} * r_{it}) \qquad \text{for all} \quad t = 1, ..., 156 \tag{3}$$

$$w_{it} = \frac{r_{it}}{\sum_{i=1}^{N} r_{it}} \qquad \text{for all} \quad t = 1, ..., 156 \tag{4}$$

$$\begin{aligned} &Where: \\ &\quad r_{pt} \;=\; Value\ weighted\ portfolio\ return\ at\ time\ t \\ &\quad r_{it} \;=\; Return\ on\ stock\ i\ at\ time\ t \end{aligned}$$

To assess what risk-free rate to use, we have to evaluate what the relevant investors' risk-free alternative is. It is reasonable to assume that the index consists of companies with several international institutional investors, which makes it hard to choose a specific risk free rate. Kenneth R. French's data library uses the US one-month T-bill rate as risk free return. However, since

[4] A zero investment strategy typically refers to forming a long portfolio in one set of securities and a short portfolio in another such that the net value is zero (Alexander, 2000).

the index consists mostly of companies within the same monetary union, we have decided to use the European equivalent, namely the ECB one-month government bond rate. As market return, we use the monthly value weighted return of the index.

4.5 The Fama French factors

The Kenneth R. French data library is an extensive database that has constructed risk factors for several developed markets (French, 2018). Through this database we obtain the risk factors for the Western European market that we use in the multi-factor asset pricing models. The markets that are used to construct these factors are the same as those represented in STOXX 600 Europe except for Poland and Luxembourg, whereas they use Greece instead (STOXX, 2018).[5]

4.6 Concerns about the dataset

Our sample excludes companies we do not have at least 13 years of continuous ESG data on. Hence, we may have excluded firms that have gone bankrupt or had too low profitability to stay in the STOXX index for the whole time span. If a significant proportion of the excluded firms are among the best or worst ESG performers, this may cause a survivorship bias. The crew at STOXX provided us with an overview of all the components of the STOXX 600 Europe index for the last five years. We used this to create a portfolio of companies that had been a part of STOXX once or more during the last five years, but not been a part of our index. With this information, we uncovered that there was a slight skewness towards low ESG performance among the excluded firms. We then calculated the monthly stock returns for the companies with a lower ESG score than the average threshold[6] for the

[5]Detailes are to be found in table 20 in the appendix

[6]By threshold, we mean the company with the highest ESG score that is still in the bad portfolio. The thresholds are 54.85 for the portfolios sorted on ESG, and 40.97 for the portfolios sorted on ESGC.

bad portfolios, and found it to be no different from the returns on the bad portfolios we use in the analysis. Thus, this gives us a robust indication that our sample is not exposed to survivorship bias.

5 Methodology

To measure the difference in performance between the good and bad portfolios, we compute the alphas of a long-short zero investment strategy, using Fama-French three-factor, four-factor (Carhart) and five-factor with and without momentum. These models are all augmentations of the CAPM, but due to the varying empirical records and the emergence of other risk factors, most applications of the model has been invalidated (Fama and French, 2003). Thus, this model will not be tested in this thesis. This section will explain the specifications of the models we are using and address the tests we have carried out to ensure the robustness of the models. We will also point out some weaknesses of the models.

5.1 Model specification

The Fama-French models try to explain the variation in stock performance of publicly traded companies (Womack and Zhang, 2003). The goal for the models is to capture all variations in stock prices, which would imply the intercept of all stocks being zero, i.e. no alpha. One of the advantages of using the Fama-French models is their underlying intuition, which is to capture risks that were empirically found to impact asset returns (Womack and Zhang, 2003). Consequently, we do not have to arrange the data to accommodate for all the different firm or industry specific risks within each portfolio. Moreover, the Fama-French risk factors are widely used by researchers and investors, and because of this recognition and model knowledge, using Fama-French will make our research more intuitive and comparable to previous research. That being said, the model output from a long-short zero investment strategy has to be interpreted differently. Since we are analyzing differences, the estimates, as well as the R-squared, may be less significant than what one would expect from a long only portfolio. If an estimate is insignificant, it means that there is no difference in exposure to this specific risk factor between the two portfolios in the long-short portfolio. For the time-series analysis, we are using least squares regressions.

5.1.1 Fama-French three-factor model

The Fama-French three-factor model expands the capabilities of CAPM by adding two company specific risk factors, SMB and HML. These two factors are used to account for the portfolio‘s exposure to size and value. Outside market risk, these two factors were found to be the most important factors for explaining publicly traded stock returns (Womack and Zhang, 2003). HML is short for “High Minus Low”, which is a portfolio that mimics a portfolio that is long high book-to-market stocks, also known as value stocks, and short low book-to-market stocks, known as growth stocks. This way, the factor accounts for a portfolio’s exposure to high value firms by measuring a value premium. SMB is short for “Small Minus Big”, which is a portfolio that mimics a portfolio that is long small cap stocks and short large cap stocks. This way, the factor accounts for a portfolio’s exposure to small cap stocks by measuring a size premium.

The model is constructed in the following way:

$$GMB_t = \alpha + \beta_{mrkt} * (mrkt_t - rf_t) + \beta_{SMB} * SMB_t + \beta_{HML} * HML_t + u_t \quad (5)$$

$$\begin{aligned}
Where: & \\
GMB_t &= Excess\ return\ on\ good\ minus\ bad\ portfolio\ at\ time\ t \\
\alpha &= Intercept/abnormal\ return \\
\beta_{mrkt} &= Exposure\ to\ the\ market\ factor\ (market\ beta) \\
mrkt_t - rf &= Excess\ return\ in\ the\ market\ at\ time\ t \\
\beta_{SMB} &= Exposure\ to\ the\ size\ factor \\
SMB_t &= The\ size\ factor\ at\ time\ t \\
\beta_{HML} &= Exposure\ to\ the\ value\ factor \\
HML_t &= The\ value\ factor\ at\ time\ t \\
u_t &= Error\ term\ at\ time\ t
\end{aligned}$$

5.1.2 Carhart four-factor model

In addition to the mentioned three factors, this model adds a factor that accounts for the persistence in performance. The factor is used to measure a portfolio's exposure to previous winners and losers, i.e. the momentum. The factor is constructed by mimicking a portfolio that is long previous winners and short losers (Carhart, 1997).

The Carhart model is constructed in the following way:

$$GMB_t = \alpha + \beta_{mrkt} * (mrkt_t - rf_t) + \beta_{SMB} * SMB_t + \beta_{HML} * HML_t + \beta_{MOM} * MOM_t + u_t \quad (6)$$

$$\begin{aligned} &Where: \\ \beta_{MOM} &= Exposure\ to\ the\ momentum\ factor \\ MOM_t &= The\ momentum\ factor\ at\ time\ t \end{aligned}$$

5.1.3 Fama-French five-factor model

The research by Fama and French has shown evidence that the five-factor model performs better than its ancestor, the three-factor model (2014b). In this model, we add the two factors, RMW and CMA. RMW stands for "Robust Minus Weak", and it represents the difference in returns between a portfolio comprising firms with robust profitability and a portfolio comprising firms with weak profitability, both diversified. CMA is short for "Conservative Minus Aggressive" in terms of investment strategy, and covers the difference in returns between a portfolio of low investment stocks and a portfolio of high investment stocks, both diversified (Fama and French, 2014a).

The five-factor model is constructed in the following way:

$$GMB_t = \alpha + \beta_{mrkt} * (mrkt_t - rf_t) + \beta_{SMB} * SMB_t + \beta_{HML} * HML_t + \beta_{RMW} * RMW_t + \beta_{CMA} * CMA_t + u_t \quad (7)$$

$$\begin{aligned} &Where: \\ \beta_{RMW} &= Exposure\ to\ the\ profitability\ factor \\ RMW_t &= The\ profitability\ factor\ at\ time\ t \\ \beta_{CMA} &= Exposure\ to\ the\ investment\ factor \\ CMA_t &= The\ investment\ factor\ at\ time\ t \end{aligned}$$

5.1.4 Fama-French five-factor model with momentum

This model is the same as the five-factor model, but with the additional momentum factor from the Carhart model. (Fama and French, 2014a).

$$GMB_t = \alpha + \beta_{mrkt} * (mrkt_t - rf_t) + \beta_{SMB} * SMB_t + \beta_{HML} * HML_t + \beta_{RMW} * RMW_t + \beta_{CMA} * CMA_t + \beta_{MOM} * MOM_t + u_t \quad (8)$$

5.2 Model testing

In order to justify the results from our analysis, we need the data to satisfy certain assumptions such as homoscedasticity and no autocorrelation[7] (Wooldrigde, 2013). Moreover, to run a time series analysis, we need to have stationary time series. If the data does not meet these requirements, we have to transform the data into doing so, or else the result may be spurious. If the data survives these tests, we can use all the features of OLS regression without restrictions.

5.2.1 Stationarity

An essential requirement for all times series data, is that the variables must be stationary (Wooldrigde, 2013). If a variable is non-stationary, it cannot

[7]Output tables for the tests are available under Model testing in the appendix

be used in a linear regression, unless it is transformed into a stationary process. To check for stationarity, we run an augmented dickey fuller test for a unit root, augmented to account for possible autocorrelation. To decide what lag length to use in the test, we use the optimal lag constructed by Ng and Perron. The tests state that the returns from the value weighted portfolios sorted on ESGC score are non-stationary, and therefore useless unless transformed. Thus, we calculate the first difference[8] of the variables, and when we test these again, they come out stationary[9].

5.3 Model weaknesses

According to Fama and French (2014b), the five-factor model performs better than the three-factor model. However, any asset pricing model is likely to be misspecified (Kapadia and Paye, 2014). One of the issues that emerge in the five-factor model is that the value factor (HML, becomes redundant when introducing the RMW and CMA factors, especially if parsimony is an issue (Fama and French, 2014a). This is because the average stock return is being completely captured by the other risk factors. Thus, if the sole interest is to estimate abnormal returns, the model performs equally well with and without the HML factor. However, Fama and French (2014b) emphasis that until further evidence is provided, the redundancy of HML could be specific to a period or a market. They also recommend using all the five factors to capture possible tilts in the portfolios.

Moreover, Fama and French (2014a) raise a concern related to controlling for additional factors, e.g. the momentum factor. When adding momentum, the correlations among the other five variables are likely to weaken the explaining power of the regression. To minimize possible misspecifications, we will not base our findings solely on one of the models. In addition, Fama and French state that the most serious problems of asset pricing models are in small stocks. The index that we are using for the analysis, is well diversi-

[8]First difference of $GMB = GMB_t - GMB_{t-1}$ for all t

[9]See appendix under Model testing for test output

fied, and compared to other markets and indices, the market cap is relatively large. Thus we believe our exposure to this model weakness is rather limited.

6 Results

This section covers the results from our analysis. We run several regressions with different dependent and independent variables in order to answer whether firms with good ESG performance perform different in the stock market than those with bad ESG performance. We are trying to explain the returns with four different asset pricing models, applying a long-short zero investment strategy which is long portfolios comprised of companies with good ESG performance and short portfolios comprised of companies with bad ESG performance. For the main analysis, we are applying monthly rebalancing portfolios, and for robustness, we will run all the models again, with yearly rebalancing portfolios.

In short, our results show some clear patterns when the portfolios are sorted on ESG scores. Nearly all the models show evidence that the bad portfolios perform better, as well as being more exposed to small cap stocks than the good portfolios. These effects are not present in the portfolios sorted on ESGC scores, which appear to be much more volatile. Each model's output and description is presented in tables on a dedicated page, followed by a written summary of the results on the next page.

6.1 Fama-French three-factor model

Table 3

Results from the Fama-French three-factor model, quartile portfolios								
	ESGC score				ESG score			
	Equally weighted		Value weighted		Equally weighted		Value weighted	
Parameter	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Intercept	0.049	0.46	−0.019	−0.22	−0.464***	−3.36	−0.489***	−2.96
MRKT-rf	−0.082**	−2.58	−0.022	−0.94	0.056	1.54	0.001	0.02
SMB	−0.051	−0.87	0.128*	1.96	−0.291***	−3.67	−0.385***	−4.04
HML	−0.055	−1.04	0.117**	2.09	0.104	1.62	0.191***	2.69
N	156		154		156		156	
R-squared(%)	15.12		5.31		12.25		14.84	
Original DW			3.030					
Transf. DW			2.276					

Results from the Fama-French three-factor model, decile portfolios								
	ESGC score				ESG score			
	Equally weighted		Value weighted		Equally weighted		Value weighted	
Parameter	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Intercept	−0.072	−0.37	0.051	0.40	−0.300**	−2.54	−0.366**	−2.39
MRKT-rf	−0.134*	−1.85	0.004	0.11	0.099***	2.70	0.049	1.15
SMB	−0.156	−1.41	−0.136	−1.29	−0.445***	−4.55	−0.491***	−4.23
HML	−0.126	−1.13	0.291***	3.70	0.066	0.82	0.017	0.20
N	156		154		155		155	
R-squared(%)	15.38		10.41		17.64		11.94	
Original DW			3.282		2.691		2.634	
Transf. DW			2.282		1.948		1.974	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table provides the results from the Fama-French three-factor model. The dependent variables are the excess return of four monthly rebalancing long-short portfolios that mimic long good ESG and short bad ESG, zero investment. The first table shows the portfolios picked on top/bottom quartiles ESG performance, and the second table shows the portfolios picked on top/bottom deciles. On the left side of the tables, the scores are based on ESGC, and on the right side, they are based on ESG. Finally, the analysis covers both equally and value weighted portfolios. The variable MRKT-rf is the value weighted market return less the risk free rate, the coefficient picks up the difference in market β between the good and the bad portfolio. The SMB factor captures the portfolios' exposure to small cap stocks, the coefficient picks up the difference in exposure between the good and the bad portfolio. The HML factor captures the portfolios' exposure to high book-to-market value firms, the coefficient picks up the difference in exposure between the good and the bad portfolio. Finally, the intercept captures the difference in abnormal return of the portfolios. We estimated the model with monthly data from 2005-2017.

In terms of abnormal return, both the large and the small portfolios show the same results. When we sort companies on ESG score, the abnormal returns among the bad stocks are on average about 0.5% higher for the large portfolios and 0.3% for the small portfolios. All four intercepts are significant. On the other hand, the portfolios based on ESGC score are not different in terms of abnormal returns, as none of the intercepts are significant.

Next, considering the systematic risk, the large, bad portfolios, sorted on ESGC score, show a higher volatility than the good portfolios, when equally weighted, but no difference when value weighed. For the small portfolios, the equally weighted, good portfolio has significantly higher volatility than the bad portfolio, when sorted on ESG score. The value weighted returns show no significant difference.

The SMB factor shows us that when we sort firms on ESGC score, there is no difference in the exposure to size between the good and the bad portfolios. When sorting on ESG score on the other hand, the bad portfolios has unambiguously higher exposure to small cap stocks than the good portfolio.

The HML factor is significant for all the value weighted portfolios except for the small portfolio sorted on ESG score. This positive exposure to HML indicates that the good portfolio comprises more high value stocks than the bad portfolio. None of the equally weighted portfolios have significant exposure to HML.

The Durbin Watson statistics on the value weighted ESGC sorted portfolios, show that the first differencing led to a strong negative autocorrelation. After using the Cochrane-Orcutt transformation, we end up with inconclusive Durbin-Watson statistics.[10][11]

[10]Meaning DW cannot conclude on netiher H_0 of no autocorrelation nor H_1 of autocorrelation(*A Durbin-Watson Significance Tables*)

[11]$d_L = 1.693, d_U = 1.774$(*A Durbin-Watson Significance Tables*)

6.2 Carhart model

Table 4

Results from the Carhart model, quartile portfolios								
	ESGC score				ESG score			
	Equally weighted		Value weighted		Equally weighted		Value weighted	
Parameter	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Intercept	−0.017	−0.13	0.009	0.09	−0.499***	−3.29	−0.480***	−2.61
MRKT-rf	−0.074***	−2.97	−0.028	−1.18	0.061*	1.70	−0.001	−0.02
SMB	−0.056	−0.89	0.132**	2.01	−0.295***	−3.68	−0.384***	−4.02
HML	0.021	0.40	0.091	1.52	0.138*	1.85	0.182**	2.43
MOM	0.082	1.34	−0.036	−0.83	0.044	0.90	−0.012	−0.18
N	155		154		156		156	
R-squared(%)	18.36		6.06		12.92		14.88	
Original DW	1.893		3.037					
Transf. DW	1.998		2.274					

Results from the Carhart model, decile portfolios								
	ESGC score				ESG score			
	Equally weighted		Value weighted		Equally weighted		Value weighted	
Parameter	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Intercept	−0.243	−0.96	0.097	0.65	−0.279**	−2.04	−0.338**	−1.99
MRKT-rf	−0.105**	−2.22	−0.006	−0.17	0.095**	2.49	0.043	1.02
SMB	−0.173	−1.45	−0.130	−1.22	−0.443***	−4.51	−0.488***	−4.20
HML	0.037	0.40	0.249***	2.94	0.047	0.52	−0.009	−0.09
MOM	0.215	1.45	−0.059	−0.88	−0.026	−0.44	−0.035	−0.58
N	156		154		155		155	
R-squared(%)	22.84		11.33		17.87		12.22	
Original DW			3.284		2.692		2.638	
Transf. DW			2.290		1.955		1.979	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table provides the results from the Carhart model. The dependent variables are the excess return of four monthly rebalancing long-short portfolios that mimic long good ESG and short bad ESG, zero investment. The first table shows the portfolios‘ picked on top/bottom quartiles ESG performance, and the second table shows the portfolios‘ picked on top/bottom deciles. On the left side of the tables, the scores are based on ESGC, and on the right side, they are based on ESG. Finally, the analysis covers both equally and value weighted portfolios. The variable MRKT-rf is the value weighted market return less the risk free rate, the SMB factor captures the portfolios‘ exposure to small cap stocks and the HML factor captures the portfolios‘ exposure to high book-to-market value firms. The MOM factor captures the portfolios’ exposure to previous winners or losers, i.e. the momentum, the coefficient picks up the difference in exposure between the good and the bad portfolio. Finally, the intercept captures the difference in abnormal return of the portfolios. We estimated the model with monthly data from 2005-2017.

Again, we are most interested in the abnormal return, which is the same for the Carhart model as for the Fama-French three-factor model. If we sort companies on ESG score, the monthly abnormal returns among the bad stocks are on average about 0.5% higher for the large portfolios and about 0.3% for the small portfolios. All four intercepts are significant. On the other hand, the portfolios based on ESGC score are not different in terms of abnormal returns, as none of the intercepts are significant.

Regarding the systematic risk, the equally weighted, bad portfolios sorted on ESGC score show higher volatility than the good portfolios. However, when using value weighted returns, these differences vanishes. For the ESG sorted portfolios, the only significant coefficient is for the small, equally weighted portfolio, whereas the coefficient states that the good portfolio is more volatile.

The SMB factor coefficients show almost the same results as for the three-factor model. The large, value weighted portfolios state that the good portfolios are more exposed to small cap stocks, when sorted on ESGC score. On the other hand, the bad portfolios have unambiguously higher exposure to small cap stocks than the good portfolios, when sorted on ESG score.

For the HML factor, we do not see any clear tendencies. Only the small, value weighted ESGC portfolio and the large, value weighted ESG portfolio are positive and significant, which means that the good portfolio is more exposed towards high value stocks.

Lastly, the Carhart model is augmented with a momentum factor, which shows no significance in any of the regressions.

Again, the Durbin-Watson statistics on the value weighted ESGC portfolios show that the first differencing led to a strong negative autocorrelation. After using the Cochrane-Orcutt transformation, we end up with inconclusive Durbin-Watson statistics.

6.3 Fama-French five-factor model

Table 5

Results from the Fama-French five-factor model, quartile portfolios

	ESGC score				ESG score			
	Equally weighted		Value weighted		Equally weighted		Value weighted	
Parameter	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Intercept	−0.025	−0.21	−0.042	−0.39	−0.481***	−3.27	−0.485***	−2.65
MRKT-rf	−0.082***	−2.78	−0.024	−0.97	0.065*	1.86	0.005	0.13
SMB	−0.026	−0.41	0.133*	1.94	−0.273***	−3.14	−0.379***	−3.70
HML	0.029	0.30	0.153	1.50	0.068	0.74	0.156	1.32
RMW	0.163	1.30	0.063	0.46	−0.011	−0.08	−0.038	−0.22
CMA	0.070	0.67	−0.014	−0.12	0.152	1.05	0.072	0.40
N	155		154		156		156	
R-squared(%)	16.28		5.49		13.42		15.06	
Original DW	1.894		2.934					
Transf. DW	2.000		2.270					

Results from the Fama-French five-factor model, decile portfolios

	ESGC score				ESG score			
	Equally weighted		Value weighted		Equally weighted		Value weighted	
Parameter	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Intercept	−0.226	−1.19	0.124	0.76	−0.321**	−2.40	−0.397**	−2.12
MRKT-rf	−0.128**	−2.06	−0.003	−0.07	0.098**	2.53	0.041	0.95
SMB	−0.104	−0.92	−0.166	−1.51	−0.439***	−4.25	−0.488***	−3.99
HML	0.023	0.10	0.228	1.51	0.100	0.83	0.093	0.54
RMW	0.343	1.52	−0.157	−0.77	0.060	0.35	0.109	0.45
CMA	0.150	0.67	−0.093	−0.49	−0.013	−0.10	−0.085	−0.49
N	156		154		155		155	
R-squared(%)	17.27		11.17		17.66		12.23	
Original DW			3.263		2.650		2.608	
Transf. DW			2.287		1.946		1.972	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table provides the results from the Fama-French five-factor model. The dependent variables are the excess return of four monthly rebalancing long-short portfolios that mimic long good ESG and short bad ESG, zero investment. The first table shows the portfolios picked on top/bottom quartiles ESG performance, and the second table shows the portfolios picked on top/bottom deciles. On the left side of the table, the scores are based on ESGC, and on the right side, they are based on ESG. Finally, the analysis covers both equally and value weighted portfolios. The variable MRKT-rf is the value weighted market return less the risk free rate, the SMB factor captures the portfolios' exposure to small cap stocks, the HML factor captures the portfolios' exposure to high book-to-market value firms. The RMW factor captures the portfolios' exposure to firms with robust profitability, the coefficient picks up the difference in exposure between the good and the bad portfolio. The CMA factor captures the portfolios' exposure to firms with a conservative investment strategy, the coefficient picks up the difference in exposure between the good and the bad portfolio. Finally, the intercept captures the difference in abnormal return of the portfolios. We estimated the model with monthly data from 2005-2017.

Based on ESG scores, this model also provides clear evidence of a higher abnormal return among the bad portfolios. Monthly about 0.5% higher for the large portfolio and 0.3-0.4% higher for the small portfolio. All four intercepts are significant. In line with the previous models, none of the portfolios based on ESGC score show any differences in abnormal returns.

Regarding the systematic risk, these results are the same as for the Carhart model. When sorting on ESGC score, the equally weighted, bad portfolios are more volatile than the good ones. For the ESG portfolios however, the only significant coefficient is for the small, equally weighted portfolio, which states that the good portfolio is more volatile. All other coefficients are insignificant.

The SMB factor coefficient is still unambiguous in sign and magnitude. There is no difference in the exposure to size between the good and the bad portfolios, when sorted on ESGC score. On the other hand, the bad portfolios have a higher exposure to small cap stocks than the good portfolios, when sorted on ESG score.

Neither the HML factor, the RMW factor, nor the CMA factor have any significant coefficients in the five-factor model.

Again, the Durbin-Watson statistics on the value weighted ESGC portfolios show that the first differencing led to a strong negative autocorrelation. After using the Cochrane-Orcutt transformation, we end up with inconclusive Durbin-Watson statistics.

6.4 Fama-French five-factor model with momentum

Table 6

Results from the Fama-French five-factor model with momentum, quartile portfolios

	ESGC score				ESG score			
	Equally weighted		Value weighted		Equally weighted		Value weighted	
Parameter	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Intercept	−0.070	−0.53	−0.016	−0.14	−0.497***	−3.16	−0.472**	−2.42
MRKT-rf	−0.076***	−3.09	−0.027	−1.08	0.067*	1.92	0.004	0.08
SMB	−0.041	−0.64	0.143**	2.07	−0.279***	−3.17	−0.375***	−3.62
HML	0.092	1.28	0.116	1.09	0.091	0.88	0.138	1.18
RMW	0.141	1.07	0.062	0.45	−0.018	−0.14	−0.032	−0.19
CMA	0.005	0.05	0.027	0.23	0.128	0.83	0.092	0.49
MOM	0.080	1.22	−0.039	−0.87	0.029	0.56	−0.022	−0.36
N	155		154		156		156	
R-squared(%)	19.15		6.25		13.67		15.16	
Original DW	1.893		2.937					
Transf. DW	1.998		2.266					

Results from the Fama-French five-factor model with momentum, decile portfolios

	ESGC score				ESG score			
	Equally weighted		Value weighted		Equally weighted		Value weighted	
Parameter	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Intercept	−0.352	−1.55	0.159	0.89	−0.303**	−2.05	−0.379*	−1.91
MRKT-rf	−0.108**	−2.41	−0.007	−0.18	0.095**	2.42	0.039	0.90
SMB	−0.148	−1.20	−0.153	−1.39	−0.434***	−4.09	−0.482***	−3.91
HML	0.177	1.42	0.176	1.09	0.075	0.59	0.068	0.36
RMW	0.280	1.18	−0.158	−0.79	0.061	0.36	0.111	0.46
CMA	−0.023	−0.15	−0.037	−0.18	0.014	0.09	−0.057	−0.31
MOM	0.214	1.39	−0.054	−0.73	−0.027	−0.43	−0.028	−0.42
N	155		154		155		155	
R-squared(%)	24.34		11.85		17.9		12.39	
Original DW	2.062		3.263		2.651		2.610	
Transf. DW	1.993		2.297		1.953		1.976	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table provides the results from the Fama-French five-factor model with momentum. The dependent variables are the excess return of four monthly rebalancing long-short portfolios that mimic long good ESG and short bad ESG. The first table shows the portfolios picked on top/bottom quartiles ESG performance, and the second table shows the portfolios picked on top/bottom deciles. On the left side of the table, the scores are based on combined score, and on the right side, they are based on ESG score. Finally, the analysis covers both equally and value weighted portfolios. The variable MRKT-rf is the value weighted market return less the risk free rate, the SMB factor captures the portfolios‘ exposure to small cap stocks, the HML factor captures the portfolios‘ exposure to high market-to-book value firms, the RMW factor captures the portfolios‘ exposure to firms with robust profitability, the CMA factor captures the portfolios‘ exposure to firms with a conservative investment strategy, and the MOM factor captures the portfolios‘ exposure to previous winners and losers. Finally, the intercept captures the difference in abnormal return of the portfolios. We estimated the model with monthly data from 2005-2017.

The Fama-French five-factor model with momentum is the last model we are running, and it does not differ much from the previous models. All the portfolios sorted on ESG score, except the small, value weighted portfolio, show higher abnormal return among the bad portfolios. About 0.5% monthly for the large portfolios and 0.3% for the small, equally weighted portfolio. The intercepts for the ESGC portfolios show no differences in the abnormal returns.

The systematic risk factor shows that when sorted on ESGC score, the equally weighted, bad portfolios are more volatile than the good portfolios. When sorted on ESG score, the small, equally weighted portfolio shows the opposite, namely that the good portfolio is more volatile.

The SMB factor coefficients follow the same pattern as for the rest of the models. The large, value weighted portfolios state that the good portfolios are more exposed to small cap stocks, when sorted on ESGC score. On the other hand, the bad portfolios have unambiguously higher exposure to small cap stocks than the good portfolios, when sorted on ESG score.

None of the remeaining variables are significant in this model.

As for the previous models, the Durbin-Watson statistics on the value weighted ESGC portfolios show that the first differencing led to a strong negative autocorrelation. After using the Cochrane-Orcutt transformation, we end up with inconclusive Durbin-Watson statistics.

6.5 Results with yearly rebalancing

In the table below, we have assembled the intercepts from the 32 regressions with yearly rebalancing portfolios[12].

Table 7

Intercepts from yearly rebalancing portfolios

	ESGC score				ESG score			
	Equally weighted		Value weighted		Equally weighted		Value weighted	
Model	Large	Small	Large	Small	Large	Small	Large	Small
3 factor	0.026	-0.175	-0.024	0.047	-0.478***	-0.383***	-0.483***	-0.413***
Carhart	-0.039	-0.349	-0.002	0.101	-0.512***	-0.366***	-0.476**	-0.397**
5 factor	-0.050	-0.301	-0.016	0.173	-0.483***	-0.439***	-0.468**	-0.480**
5 factor+mom	-0.096	-0.432**	0.005	0.209	-0.502***	-0.427***	-0.459**	-0.471**

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table presents the intercepts from the regressions with yearly rebalancing portfolios.

As we can see from the table, the sign and magnitude of the intercepts do not weaken when we rebalance the portfolios yearly instead of monthly. When we sort companies on ESG score, the monthly abnormal returns are still about 0.5% higher for the large, bad portfolios. The difference is smaller and more varying for the small portfolios, but still significant. As for the portfolios sorted on ESGC score, the results are still mostly insignificant. The only exception is for the small, equally weighted portfolio estimated by the five-factor model with momentum, whereas the bad portfolio again outperforms the good.

[12]The full regression output can be found under Output tables, yearly rebalancing in the appendix

7 Discussion

This thesis attempts to answer whether companies with good ESG performance perform different in the stock market than those with bad ESG performance. We study a selection of 401 companies from the European index, STOXX Europe 600, and from this selection, we have sorted out eight portfolios, four monthly rebalanced on the basis of ESG score and four rebalanced on the basis of ESGC score[13]. To measure the difference in stock performance, we compute the alphas of a long-short zero investment strategy which is long the good portfolios and short the bad portfolios. We apply Fama-French three-factor, four-factor (Carhart) and five-factor, with and without momentum to account for potential differences in risk exposure between the portfolios. In this section, we will discuss our most prominent findings, and in the following table we have assembled the results from all the models with monthly rebalancing portfolios.

Table 8

Overview all models								
	ESGC				ESG			
	Equally weighted		Value weighted		Equally weighted		Value weighted	
Parameter	Large	Small	Large	Small	Large	Small	Large	Small
Intercept	nnnn	nnnn	nnnn	nnnn	wwww	wwww	wwww	wwwn
MRKT-rf	wwww	nwww	nnnn	nnnn	nnnn	bbbb	nnnn	nnnn
SMB	nnnn	nnnn	nbnb	nnnn	wwww	wwww	wwww	wwww
HML	nnnn	nnnn	bnnn	bbnn	nnnn	nnnn	bbnn	nnnn
RMW	- -nn	- -nn	- -nn	- -nn	- -nn	- -nn	- -nn	- -nn
CMA	- -nn	- -nn	- -nn	- -nn	- -nn	- -nn	- -nn	- -nn
MOM	-n-n	-n-n	-n-n	-n-n	-n-n	-n-n	-n-n	-n-n

In this table we have collected all the results from the models based on monthly rebalancing portfolios, to provide a better visual overview of the results from the regressions. n = not significant (95%), b = good is better, w = good is worse.* The letters are sorted in the way the models are presented in the results section, thus the leftmost letter is the Fama-French three-factor model, followed by the Carhart model, the five-factor model and finally the five-factor model with momentum.

*Note that the exposures to risk factors in asset pricing models are meant to explain the variation in stock returns, and that positive(negative) coefficients do not necessarily mean better(worse), however, we needed an intuitive and lucid way to present the results.

[13]Explanation for the ESG and the ESGC score is given in detail under the data section.

7.1 Abnormal returns

We have shown that portfolios comprising stocks with low ESG score outperform portfolios comprising stocks with high ESG score. All portfolios give the same unequivocal evidence. For the small portfolios the monthly difference is on average around 0.3-0.4% and for the large portfolios on average around 0.5%.

The explanation for these differences may be framed in terms of those processes and practices differentiating the two groups. In other words, those practices that companies with high ESG score carry out to achieve a high rating, and companies with a low score do not. Some of the practices such as stakeholder engagement and measuring and reporting of nonfinancial performance metrics are either added costs, or resources that could be spent differently (Eccles et al. 2014). These costs may outweigh the benefit provided by being socially responsible, such as attracting better human capital, or avoiding costly controversies. Furthermore, they may prevent an efficient resource allocation, by making the company forgo valuable business opportunities that do not fit their sustainability profile.

The findings may also be explained by management focus (Eccles et al. 2014). Managements and boards of high scoring ESG companies may be distracted from the business' main goal, increasing the profit. They may be focused on sustainability rather than creating shareholder value, and perhaps confusing the market as well. Time spent on sustainability issues is time that could be spent on other key issues, and therefore indirectly destroy shareholder value. Another explanation for the good portfolios' weaker stock performance may be that they are increasing the risk, through creating expectations that cannot be met. Satisfying stakeholders, might cause them to raise their demands, which further can weaken a company's ability to provide value to them.

Moreover, our findings challenge the arguments made by several articles in the media lately. Based on resource slack theory, they have argued that good

profitability allows companies to invest in ESG and therefore should expect to see a link between ESG and CFP (e.g. Finansavisen, 2018). If this statement is correct, we should not see a significantly lower stock performance for companies with high ESG scores. To back it up, there is no impact on the RMW risk factor, which indicates that there are no differences in the robustness of the profitability. A counterargument can be made upon the fact that it is costly to develop ESG and to improve an ESG score. However, since the scores are calculated using a percentile rank methodology, it is reasonable to assume that staying among the best requires continuous improvements and costs.

Interestingly, the findings disappear completely if we account for ESG controversies, i.e. when the portfolios are sorted on the basis of ESGC score. We can explain this by the portfolios sorted on ESGC being more exposed to systematic risk, when equally weighted. This makes sense in the way that those companies that most frequently get media attention fluctuate more than the rest. We have seen that the stock composition in the portfolios sorted on ESGC scores differs quite a lot from the ones sorted on ESG scores. There is a higher turnover of stocks coming in and out of the portfolios as they rebalance every month, which further indicates that the ESGC score is more volatile. Since it does not follow the ESG scores, it might stipulate that all companies are affected by controversies independently of their internal ESG strategy.

7.2 Exposure to size

We also see a large significant difference in size when we sort the portfolios on ESG score, i.e. the companies in the good portfolios are larger than the companies in the bad portfolios. This suggests that it is the large companies that embrace ESG, and not necessarily the most profitable. The findings are also in line with previous research (e.g. Vermeir et al., 2005) and can be explained with the argument that companies that manage values with a high societal interest are consequently under more pressure from government,

media and society expectations. A simultaneous effect of the media's focus on these companies is that they also get the most attention when something is wrong. Hence, there may be a dissonance between the actual ESG commitment of a company, and the amount of negative attention the company receives in the media, which also can explain why the difference in size is wiped out when companies are sorted on ESGC score.

7.3 Implications

Our findings imply that it can be costly to bet on ESG, and for investors exclusively seeking shareholder value, a profitable strategy can oppositely be to invest in companies with low ESG score. On the other hand, the trading costs of such a strategy might absorb the profits. For this purpose, the yearly rebalancing portfolios might illustrate a more realistic trading strategy,[14] even though also this will generate trading costs as well as holding costs for the short positions. To formulate a strategy on the basis of the ESGC score is harder since ESG controversies seem to hurt stocks more randomly. With that being said, this research must not be mistaken for a concrete trading recommendation. We apply a long-short zero investment strategy to asses the difference in stock performance between the good and bad ESG performers and address ESG as a possible risk factor.

[14]The intercepts are shown in table 7 under the results section, full output tables are available under Output tables, yearly rebalancing in the appendix.

8 Conclusion

The main purpose of this study is to compare the difference in stock performance between companies with good and bad ESG performance. Our results produce evident material that stocks with low ESG score outperform stocks with high ESG score. The finding contradicts our hypothesis as well as the majority of previous research on sustainable stock performance. Our results are relevant for anyone seeking a sustainable investment, especially in Europe, and imply that it can be costly to bet on ESG. For investors exclusively seeking shareholder value, a profitable strategy may oppositely be to bet against ESG. The question arising from these findings is whether the results are robust for alternative measurements, for example based on assessments from other rating agencies, or through alternative ways of measuring financial performance, e.g. through accounting measures.

Moreover, we find that the differences disappear as soon as we account for ESG controversies, i.e. compare the differences in stock performance between companies with high and low ESGC scores. Not only does the ESGC score exhibit a much more volatile behaviour than the ESG score, which makes it less useful for investors, but our discussion presents arguments for the dissonance between the actual ESG commitment of a company and the amount of negative attention from media and society.

Due to the difference in findings associated with the different scores, it is difficult to make any clear conclusion. What becomes clear however, is that we do not see a positive relationship between ESG and stock performance for the STOXX 600 Europe, using Thomson Reuters ESG Scores. Moreover, we believe that ESG represents an unaccounted for risk, and that the GMB factor belongs in future asset pricing models.

References

Alexander, Gordon J. (2000). "On Back-Testing "Zero-Investment" Strategies". In: *The Journal of Business* 73.2, pp. 255–278. DOI: 10.1086/209642. URL: https://www.jstor.org/stable/10.1086/209642.

Alexander, Gordon J. and Rogene A. Buchholz (1978). "Corporate Social Responsibility and Stock Market Performance". In: *Academy of Management Journal* 21.3, pp. 479–486. DOI: 10.5465/255728. URL: http://journals.aom.org/doi/10.5465/255728.

Aupperle, K. E., A. B. Carroll, and J. D. Hatfield (1985). "An Empirical Examination of the Relationship Between Corporate Social Responsibility and Profitability". In: *Academy of Management Journal* 28.2, pp. 446–463. DOI: 10.2307/256210. URL: http://amj.aom.org/cgi/doi/10.2307/256210.

Bauer, Rob, Jeroen Derwall, and Rogér Otten (2007). "The Ethical Mutual Fund Performance Debate: New Evidence from Canada". In: *Journal of Business Ethics* 70.2, pp. 111–124. DOI: 10.1007/s10551-006-9099-0. URL: http://link.springer.com/10.1007/s10551-006-9099-0.

Bauer, Rob, Kees Koedijk, and Rogér Otten (2005). "International evidence on ethical mutual fund performance and investment style". In: *Journal of Banking & Finance* 29, pp. 1751–1767. DOI: 10.1016/j.jbankfin.2004.06.035. URL: www.elsevier.com.

Bauer, Rob, Rogér Otten, and Alireza Tourani Rad (2006). "Ethical investing in Australia: Is there a financial penalty?" In: DOI: 10.1016/j.pacfin.2004.12.004. URL: www.elsevier.com/locate/pacfin.

Cai, Ye, Hoje Jo, and Carrie Pan (2012). "Doing Well While Doing Bad? CSR in Controversial Industry Sectors". In: *Journal of Business Ethics* 108.4, pp. 467–480. DOI: 10.1007/s10551-011-1103-7. URL: http://link.springer.com/10.1007/s10551-011-1103-7.

Carhart, Mark M (1997). *On Persistence in Mutual Fund Performance*. Tech. rep. 1, pp. 57–82. URL: http://personal.psu.edu/qxc2/fin597/5-carhart-JF.pdf.

Cheng, Beiting, Ioannis Ioannou, and George Serafeim (2014). "Corporate social responsibility and access to finance". In: *Strategic Management Journal* 35.1, pp. 1–23. DOI: 10.1002/smj.2131. URL: http://doi.wiley.com/10.1002/smj.2131.

Douglas, Elyse, Tracy Van Holt, and Whelan Tenise (2017). *Responsible Investing: Guide to ESG Data Providers and Relevant Trends*. Tech. rep. URL: https://cbey.yale.edu/sites/default/files/Responsible%20Investing%20-%20Guide%20to%20ESG%20Data%20Providers%20and%20Relevant%20Trends.pdf?fbclid=IwAR0hHX5rdApckPrVjxXRK4SoqxhPj-sbt9hCxLpVOsPf3ujfqyIu_wsU2pE.

Durbin, James and Geoffrey Watson. *A Durbin-Watson Significance Tables*. Tech. rep. URL: https://www3.nd.edu/~wevans1/econ30331/Durbin_Watson_tables.pdf.

Eccles, Robert G, Ioannis Ioannou, and George Serafeim (2014). "The Impact of Corporate Sustainability on Organizational Processes and Performance". In: *Management Science* 60.11, pp. 2835–2857. DOI: 10.1287/mnsc.2014.1984. URL: http://pubsonline.informs.org.https://doi.org/10.1287/mnsc.2014.1984http://www.informs.org.

Eccles, Robert G. and Michael P. Krzus (2010). *One report : integrated reporting for a sustainable strategy*. John Wiley & Sons, p. 235. ISBN: 9780470587515. URL: https://www.wiley.com/en-us/One+Report%3A+Integrated+Reporting+for+a+Sustainable+Strategy-p-9780470587515.

FAMA, EUGENE F. and KENNETH R. FRENCH (1992). "The Cross-Section of Expected Stock Returns". In: *The Journal of Finance* 47.2, pp. 427–465. DOI: 10.1111/j.1540-6261.1992.tb04398.x. URL: http://doi.wiley.com/10.1111/j.1540-6261.1992.tb04398.x.

Fama, Eugene F and Kenneth R French (1993). "Common risk factors in the returns on stocks and bonds*". In: *Journal of Financial Economics* 33, pp. 3–56. URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.139.5892&rep=rep1&type=pdf.

Fama, Eugene F. and Kenneth R. French (1998). "Value versus Growth: The International Evidence". In: *The Journal of Finance* 53.6, pp. 1975–1999.

DOI: 10.1111/0022-1082.00080. URL: http://doi.wiley.com/10.1111/0022-1082.00080.

Fama, Eugene F. and Kenneth R. French (2003). "The Capital Asset Pricing Model: Theory and Evidence". In: *CRSP Working Paper* No. 550. DOI: 10.2139/ssrn.440920. URL: http://www.ssrn.com/abstract=440920.

— (2014a). "A Five-Factor Asset Pricing Model". In: *Fama-Miller Working Paper*. DOI: 10.2139/ssrn.2287202. URL: http://www.ssrn.com/abstract=2287202.

— (2014b). *What does it mean to say HML is redundant*. Tech. rep. Fama/French Forum. URL: https://famafrench.dimensional.com/questions-answers/qa-what-does-it-mean-to-say-hml-is-redundant.aspx.

Fama, Eugene F et al. (1995). "Size and Book-to-Market Factors in Earnings and Returns". In: URL: https://www.ivey.uwo.ca/cmsmedia/3775516/size_and_book-to-market_factors.pdf.

Fauzi, Hasan and Kamil M. Idris (2009). "The Relationship of CSR and Financial Performance: New Evidence from Indonesian Companies". In: *Issues In Social And Environmental Accounting* 3.1, p. 66. DOI: 10.22164/isea.v3i1.38. URL: http://isea.icseard.uns.ac.id/index.php/isea/article/view/38.

Finansavisen (2018). *Bærekraft - Et overskuddsfenomen*.

French, Kenneth R. (2018). *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.

Friedman, Milton (1970). "The Social Responsibility of Business is to Increase its Profits The New York Times Magazine". In: URL: http://umich.edu/~thecore/doc/Friedman.pdf.

— (1982). *Capitalism and freedom*. University of Chicago Press, p. 202. ISBN: 9780226264189.

Galant, Adriana and Simon Cadez (2017). "Corporate social responsibility and financial performance relationship: a review of measurement approaches". In: *Economic Research-Ekonomska Istraživanja* 30.1, pp. 676–693. DOI: 10.1080/1331677X.2017.1313122. URL: https://www.tandfonline.com/doi/full/10.1080/1331677X.2017.1313122.

Geraghty, Michael and Sebastian Vanderzeil (2017). *The State of the Data: Challenges to Integrating ESG Factors Into Investment Decisions*. URL: https://cornerstonecapinc.com/the-state-of-the-data-challenges-to-integrating-esg-factors-into-investment-decisions/.

Global Sustainable Investment Alliance (2016). *Global Sustainable Investment Review*. Tech. rep. URL: http://www.gsi-alliance.org/wp-content/uploads/2017/03/GSIR_Review2016.F.pdf.

Griffin, Jennifer J. and John F. Mahon (1997). "The Corporate Social Performance and Corporate Financial Performance Debate". In: *Business & Society* 36.1, pp. 5–31. DOI: 10.1177/000765039703600102. URL: http://journals.sagepub.com/doi/10.1177/000765039703600102.

Hamilton, Sally, Hoje Jo, and Meir Statman (1993). "Doing Well While Doing Good? The Investment Performance of Socially Responsible Mutual Funds". In: *Financial Analysts Journal* 49.6, pp. 62–66. DOI: 10.2469/faj.v49.n6.62. URL: http://www.cfapubs.org/doi/full/10.2469/faj.v49.n6.62.

Hong, Harrison and Marcin Kacperczyk (2009). "The price of sin: The effects of social norms on markets $". In: *Journal of Financial Economics* 93, pp. 15–36. DOI: 10.1016/j.jfineco.2008.09.001. URL: www.elsevier.com/locate/jfec.

Jensen, Michael C. (1968). "THE PERFORMANCE OF MUTUAL FUNDS IN THE PERIOD 1945-1964". In: *The Journal of Finance* 23.2, pp. 389–416. DOI: 10.1111/j.1540-6261.1968.tb00815.x. URL: http://doi.wiley.com/10.1111/j.1540-6261.1968.tb00815.x.

Kapadia, Nishad and Bradley S. Paye (2014). "Estimating the Cost of Equity: Why Do Simple Benchmarks Outperform Factor Models?" In: *SSRN Electronic Journal*. DOI: 10.2139/ssrn.2118920. URL: http://www.ssrn.com/abstract=2118920.

Kell, George (2018). *The Remarkable Rise Of ESG*. URL: https://www.forbes.com/sites/georgkell/2018/07/11/the-remarkable-rise-of-esg/#65592c1e1695.

Kempf, Alexander and Peer Osthoff (2007). "The Effect of Socially Responsible Investing on Portfolio Performance". In: *European Financial Man-*

agement 13.5, pp. 908–922. DOI: 10.1111/j.1468-036X.2007.00402.x. URL: http://doi.wiley.com/10.1111/j.1468-036X.2007.00402.x.

Lima, Pedro (2017). *Fama-French Models Application to the analysis of FTSE4GOOD, MSCI ESG and STOXX ESG Indices: Do SRI Indices have biases?* Tech. rep. URL: https://repositorio-aberto.up.pt/bitstream/10216/108203/2/224407.pdf.

Mallin, C.A., B. Saadouni, and R.J. Briston (1995). "The FinancialPerformance Of Ethical Investment Funds". In: *Journal of Business Finance & Accounting* 22.4, pp. 483–496. DOI: 10.1111/j.1468-5957.1995.tb00373.x. URL: http://doi.wiley.com/10.1111/j.1468-5957.1995.tb00373.x.

McIntosh, Graham and Janusz Brzeszczyński (2014). "Performance of Portfolios Composed of British SRI Stocks". In: *Journal of Business Ethics* 120.3, pp. 335–362. DOI: 10.1007/s10551-012-1541-x. URL: http://link.springer.com/10.1007/s10551-012-1541-x.

Miles, Morgan. P. and J. Covin (2000). "Environmental Marketing: A Source of Reputational, Competitive, and Financial Advantage". In: URL: https://www.researchgate.net/publication/275971488_Miles_M_Covin_J_2000_Environmental_Marketing_A_Source_of_Reputational_Competitive_and_Financial_Advantage_Journal_of_Business_Ethics_233_299-311.

Orlitzky, Marc, Frank L. Schmidt, and Sara L. Rynes (2003). "Corporate Social and Financial Performance: A Meta-Analysis". In: *Organization Studies* 24.3, pp. 403–441. DOI: 10.1177/0170840603024003910. URL: http://journals.sagepub.com/doi/10.1177/0170840603024003910.

Pava, Moses L. and Joshua Krausz (1996). "The association between corporate social-responsibility and financial performance: The paradox of social cost". In: *Journal of Business Ethics* 15.3, pp. 321–357. DOI: 10.1007/BF00382958. URL: http://link.springer.com/10.1007/BF00382958.

Peng, Chih-Wei and Mei-Ling Yang (2014). "The Effect of Corporate Social Performance on Financial Performance: The Moderating Effect of Ownership Concentration". In: *Journal of Business Ethics* 123.1, pp. 171–182.

DOI: 10.1007/s10551-013-1809-9. URL: http://link.springer.com/10.1007/s10551-013-1809-9.

Polk, Davis and Wardwell Llp (2017). *ESG Reports and Ratings: What They Are, Why They Matter?* Tech. rep. DavisPolk. URL: https://www.davispolk.com/files/2017-07-12_esg_reports_ratings_what_they_are_why_they_matter_0.pdf.

RBC Global Asset Management (2017). *Corporate Governance & Responsible Investment*. Tech. rep. URL: http://www.rbcgam.com/corporate-governance-and-responsible-investment/pdf/2017-annual-report.pdf.

Renneboog, Luc, Jenke Ter Horst, and Chendi Zhang (2008). "Socially responsible investments: Institutional aspects, performance, and investor behavior". In: *Journal of Banking & Finance* 32.9, pp. 1723–1742. DOI: 10.1016/j.jbankfin.2007.12.039. URL: http://linkinghub.elsevier.com/retrieve/pii/S0378426607004220.

Reuters, Thomson (2018a). *Datastream Macroeconomic Analysis*. URL: https://www.refinitiv.com/en/products/datastream-macroeconomic-analysis.

— (2018b). *Thomson Reuters Datastream*.

— (2018c). *Thomson Reuters ESG Scores*. Tech. rep. URL: https://www.refinitiv.com/content/dam/gl/en/documents/methodology/esg-scores-methodology.pdf.

Roberts, Robin W. (1992). "Determinants of corporate social responsibility disclosure: An application of stakeholder theory". In: *Accounting, Organizations and Society* 17.6, pp. 595–612. DOI: 10.1016/0361-3682(92)90015-K. URL: https://www.sciencedirect.com/science/article/pii/036136829290015K.

Sharpe, William F. (1964). "Capital Asset Prices: A Theory Of Market Equilibrium Under Conditions Of Risk". In: *The Journal of Finance* 19.3, pp. 425–442. DOI: 10.1111/j.1540-6261.1964.tb02865.x. URL: http://doi.wiley.com/10.1111/j.1540-6261.1964.tb02865.x.

Sharpe, William F (1966). *Mutual Fund Performance*. Tech. rep. 1, pp. 119–138. URL: http://www.stat.ucla.edu/~nchristo/statistics_c183_c283/sharpe__mutual_fund_performance.pdf.

Statman, Meir (2000). "Socially Responsible Mutual Funds (corrected)". In: *Financial Analysts Journal* 56.3, pp. 30–39. DOI: 10.2469/faj.v56.n3.2358. URL: http://www.cfapubs.org/doi/full/10.2469/faj.v56.n3.2358.

Statman, Meir and Denys Glushkov (2008). "The Wages of Social Responsibility". In: *SSRN Electronic Journal*. DOI: 10.2139/ssrn.1372848. URL: http://www.ssrn.com/abstract=1372848.

STOXX (2018). *STOXX Digital — STOXX® Europe 600*. URL: https://www.stoxx.com/index-details?symbol=SXXP.

The Forum for Sustainable and Responsible Investment (2016). *Annual Report*. Tech. rep. URL: https://www.ussif.org/files/Publications/2016USSIFAnnualReport_online.pdf.

Treynor, Jack (1965). *Harvard business review*. Graduate School of Business Administration, Harvard University. URL: http://www.worldcat.org/title/how-to-rate-management-of-investment-funds/oclc/892276842&referer=brief_results.

Turker, Duygu (2009). "Measuring Corporate Social Responsibility: A Scale Development Study". In: *Journal of Business Ethics* 85.4, pp. 411–427. DOI: 10.1007/s10551-008-9780-6. URL: http://link.springer.com/10.1007/s10551-008-9780-6.

Ullmann, Arieh A. (1985). "Data in Search of a Theory: A Critical Examination of the Relationships Among Social Performance, Social Disclosure, and Economic Performance of U.S. Firms". In: *Academy of Management Review* 10.3, pp. 540–557. DOI: 10.5465/amr.1985.4278989. URL: http://journals.aom.org/doi/10.5465/amr.1985.4278989.

Vance, Stanley C. (1975). "Are Socially Responsible Corporations Good Investment Risks?" In:

Vermeir, Wim, Filip Corten, and Eveline Van de Velde (2005). "Corporate social responsibility and financial performance". In: *Corporate Governance: The international journal of business in society* 5.3. Ed. by Lenssen,

Gilbert, pp. 129–138. DOI: 10.1108/14720700510604760. URL: http://www.emeraldinsight.com/doi/10.1108/14720700510604760.

Waddock, Sandra and Samuel Graves (1997). "The Corporate Social Perofrmance-Financial Performance Link". In: *Strategic Management Journal* 18.4, pp. 303–319. DOI: 10.1002/(SICI)1097-0266(199704)18:4<303::AID-SMJ869>3.0.CO;2-G. URL: http://doi.wiley.com/10.1002/%28SICI%291097-0266%28199704%2918%3A4%3C303%3A%3AAID-SMJ869%3E3.0.CO%3B2-G.

Womack, Kent L. and Ying Zhang (2003). "Understanding Risk and Return, the CAPM, and the Fama-French Three-Factor Model". In: *Tuck Case* No. 03-111. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=481881.

Wooldrigde, Jeffrey M. (2013). *Introductory Econometrics - A Modern Approach*. 5th. South-Western, Cengage Learning.

9 Appendix

9.1 Model testing

Breusch-Pagan test for homoscedasticity

Table 9

Test for Homoscedasticity large portfolios								
	ESGC				ESG			
	Equally weighted		Value weighted		Equally weighted		Value weighted	
Models	F-stat	P-value	F-stat	P-value	F-stat	P-value	F	P-value
Three-factor	7.73	0.000	2.34	0.076	4.38	0.006	5.74	0.001
Carhart	7.71	0.000	0.99	0.414	4.34	0.002	5.93	0.000
Five-factor	5.63	0.000	2.14	0.064	2.34	0.044	3.21	0.009
Five-factor + mom	6.13	0.000	1.58	0.156	3.06	0.008	3.95	0.001

This table shows the results of the Breusch-Pagan test for heteroscedasticity, covering the portfolios constructed by the quartile top and bottom ESG-scores. H_0 for this test is homoscedasticity. Hence, a high P-value means we do not have a problem. As comes out clear from the table, it is safe to say that most of the regressions has a heteroscedasticity problem. Accordingly we will run the regression with robust standard errors.

Table 10

Test for homoscedasticity small portfolios								
	ESGC				ESG			
	Equally weighted		Value weighted		Equally weighted		Value weighted	
Models	F-stat	P-value	F-stat	P-value	F-stat	P-value	F-stat	P-value
Three-factor	4.93	0.003	0.87	0.459	0.85	0.469	2.16	0.095
Carhart	10.73	0.000	1.32	0.264	2.46	0.048	2.61	0.038
Five-factor	4.35	0.001	0.73	0.602	0.54	0.744	1.01	0.414
Five-factor + mom	8.62	0.000	1.00	0.424	1.60	0.151	1.48	0.188

This table shows the results of the Breusch-Pagan test for heteroscedasticity, covering the portfolios constructed by the 10th percentile top and bottom ESG-scores. H_0 for this test is homoscedasticity. Hence, a high P-value means we do not have a problem. As comes out clear from the table, it is safe to say that most of the regressions has a heteroscedasticity problem. Accordingly we will run the regression with robust standard errors.

Breusch-Godfrey test for autocorrelation

Table 11

Test for Autocorrelation large portfolios								
	ESGC				ESG			
	Equally weighted		Value weighted		Equally weighted		Value weighted	
Models	Chi2	P-value	Chi2	P-value	Chi2	P-value	Chi2	P-value
Three-factor	20.086	0.066	52.186	0.000	15.570	0.212	18.559	0.100
Carhart	21.412	0.045	53.675	0.000	15.972	0.193	18.543	0.100
Five-factor	24.423	0.018	41.448	0.000	16.428	0.172	19.358	0.080
Five-factor + mom	24.967	0.015	41.421	0.000	16.621	0.164	19.375	0.080

This table shows the results from the Breauch-Godfrey test for autocorrelation, covering the large portfolios. H_0 for this test is that there is no autocorrelation present in the regression. Hence, if we obtain a large Chi-value and low P-value, we have a problem. In this test it is very clear that many of the regressions have a problem with autocorrelation, hence we will use Cochrane-Orcutt estimates for these regressions. Under the value weighted ESGC portfolios, we can see that taking the first difference of the dependent variable has led to strong negative autocorrelation. Since we use monthly data, we have used 12 lags for the test.

Table 12

Test for Autocorrelation small portfolios								
	ESGC				ESG			
	Equally weighted		Value weighted		Equally weighted		Value weighted	
Models	Chi2	P-value	Chi2	P-value	Chi2	P-value	Chi2	P-value
Three-factor	12.693	0.392	77.136	0.000	32.734	0.001	33.358	0.001
Carhart	18.865	0.092	79.211	0.000	32.776	0.001	33.398	0.001
Five-factor	16.277	0.179	76.534	0.000	30.990	0.002	31.279	0.002
Five-factor + mom	24.875	0.015	77.763	0.000	31.004	0.002	31.244	0.002

This table shows the results from the Breauch-Godfrey test for autocorrelation, covering the small portfolios. H_0 for this test is that there is no autocorrelation present in the regression. Hence, if we obtain a large Chi-value and low P-value, we have a problem. In this test it is very clear that most of the regressions have a problem with autocorrelation, hence we will use Cochrane-Orcutt estimates for these regressions. Under the value weighted ESGC portfolios, we can see that taking the first difference of the dependent variable has led to strong negative autocorrelation. Since we use monthly data, we have used 12 lags for the test.

Augmented Dickey-Fuller test for stationarity

Table 13

Test for stationarity								
	ESGC				ESG			
	Equally weighted		Value weighted		Equally weighted		Value weighted	
Parameters	Z (t)	P-value	Z (t)	P-value	Z (t)	P-value	Z (t)	P-value
GMB large	-4.073	0.001	-2.705	0.073	-13.577	0.000	-3.261	0.017
GMB small	-5.955	0.000	-2.835	0.053	-2.934	0.042	-3.732	0.004
GMB large 1. diff			-5.475	0.000				
GMB small 1. diff			-5.582	0.000				

Parameters	Z (t)	P-value
MRKT-RF	-3.457	0.009
SMB	-3.603	0.006
HML	-2.919	0.043
RMW	-6.013	0.000
CMA	-4.051	0.001
MOM	-4.960	0.000

This table shows the results of the Augmented Dickey-Fuller test for Stationarity for all the dependent and independent variables used in the thesis. The H_0 for this test is that there is a unit root, i.e. non-stationarity. Hence, a small $|Z(t)|$ means we have a problem. As shown in the table, the return of both long-short portfolios under the value weighted ESGC sorting, fall under the H_0 of a unit root, i.e. non-stationarity. We have transformed the variables by taking the first difference, and according to the the clear rejection of H_0, the variables are now stationary and can be used in OLS regression.

9.2 ESG controversies measures

Table 14

List of all controversy measures that make up the ESG Controversy Category Score		
Category	Label (l)	Description (d)
Community	Anti-Competition Controversy	Number of controversies published in the media linked to anti-competitive behavior (e.g., antitrust and monopoly), price-fixing or kickbacks.
Community	Business Ethics Controversies	Number of controversies published in the media linked to business ethics in general, political contributions or bribery and corruption.
Community	Intellectual Property Controversies	Number of controversies published in the media linked to patents and intellectual property infringements.
Community	Critical Countries Controversies	Number of controversies published in the media linked to activities in critical, undemocratic countries that do not respect fundamental human rights principles.
Community	Public Health Controversies	Number of controversies published in the media linked to public health or industrial accidents harming the health and safety of third parties (non-employees and non-customers).
Community	Tax Fraud Controversies	Number of controversies published in the media linked to tax fraud, parallel imports or money laundering.
Human Rights	Child Labor Controversies	Number of controversies published in the media linked to use of child labor issues.
Human Rights	Human Rights Controversies	Number of controversies published in the media linked to human rights issues.
Management	Mgt Compensation Controversies Count	Number of controversies published in the media linked to high executive or board compensation.
Product Responsibility	Consumer Controversies	Number of controversies published in the media linked to consumer complaints or dissatisfaction directly linked to the company's products or services.
Product Responsibility	Controversies Customer Health & Safety	Number of controversies published in the media linked to customer health & safety.
Product Responsibility	Controversies Privacy	Number of controversies published in the media linked to employee or customer privacy and integrity.
Product Responsibility	Controversies Product Access	Number of controversies published in the media linked to product access.
Product Responsibility	Controversies Responsible Marketing	Number of controversies published in the media linked to the company's marketing practices, such as over-marketing of unhealthy food to vulnerable consumers.
Product Responsibility	Controversies Responsible R&D	Number of controversies published in the media linked to responsible R&D.
Resource Use	Environmental Controversies	Number of controversies related to the environmental impact of the company's operations on natural resources or local communities.
Shareholders	Accounting Controversies Count	Number of controversies published in the media linked to aggressive or non-transparent accounting issues.
Shareholders	Insider Dealings Controversies Count	Number of controversies published in the media linked to insider dealings and other share price manipulations.
Shareholders	Shareholder Rights Controversies Count	Number of controversies linked to shareholder rights infringements published in the media
Workforce	Diversity and Opportunity Controversies	Number of controversies published in the media linked to workforce diversity and opportunity (e.g., wages, promotion, discrimination and harassment).
Workforce	Employee Health & Safety Controversies	Number of controversies published in the media linked to workforce health and safety.
Workforce	Condition Wages or Working Condition Controversies Count	Number of controversies published in the media linked to the company's relations with employees or relating to wages or wage disputes.
Workforce	Management Departures	Has an important executive management team member or a key team member announced a voluntary departure (other than for retirement) or been ousted?

9.3 Output tables, yearly rebalancing

Fama-French three-factor model

Table 15

Results from the Fama-French three-factor model, quartile portfolios

	ESGC score				ESG score			
	Equally weighted		Value weighted		Equally weighted		Value weighted	
Parameter	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Intercept	0.026	0.25	−0.024	−0.28	−0.478***	−3.43	−0.483***	−2.86
MRKT-rf	−0.071**	−2.17	−0.014	−0.64	0.050	1.37	−0.011	−0.25
SMB	−0.061	−1.08	0.131**	2.01	−0.278***	−3.34	−0.374***	−3.84
HML	−0.076	−1.42	0.072	1.35	0.108*	1.66	0.189***	2.67
N	156		154		156		156	
R-squared (%)	14.03		3.78		11.37		14.16	
Original DW			3.059					
Transf. DW			2.331					

Results from the Fama-French three-factor model, decile portfolios

	ESGC score				ESG score			
	Equally weighted		Value weighted		Equally weighted		Value weighted	
Parameter	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Intercept	−0.175	−0.92	0.047	0.37	−0.383***	−3.35	−0.413***	−2.64
MRKT-rf	−0.117	−1.61	0.011	0.28	0.100***	2.73	0.042	0.96
SMB	−0.120	−1.09	−0.131	−1.27	−0.378***	−4.06	−0.451***	−3.93
HML	−0.153	−1.36	0.310***	3.79	0.091	1.21	0.013	0.15
N	156		154		155		155	
R-squared (%)	13.58		11.8		16.9		10.23	
Original DW			3.301		2.701		2.573	
Transf. DW			2.322		1.932		1.985	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table provides the results from the Fama-French Three-factor model. The dependent variables are the excess return of four yearly rebalancing long-short portolios that mimics long good ESG and short bad ESG, zero investment. The first table shows the portfolios picked on top/bottom quartiles ESG performance, and the second table shows the portfolios picked on top/bottom deciles. On the left side of the tables, the scores are based on ESGC, and on the right side, they are based on ESG. Finally, the analysis covers both equally and value weighted portfolios. The variable MRKT-rf is the value weighted market return less the risk free rate, the coefficient picks up the difference in market β between the good and the bad portfolio. The SMB factor captures the portfolios exposure to small cap stocks, the coefficient picks up the difference in exposure between the good and the bad portfolio. The HML factor captures the portfolios exposure to high book-to-market value firms, the coefficient picks up the difference in exposure between the good and the bad portfolio. Finally, the intercept captures the difference in abnormal return of the portfolios. We estimated the model with monthly data from 2005-2017.

Carhart model

Table 16

Results from the Carhart model, quartile portfolios

	ESGC score				ESG score			
	Equally weighted		Value weighted		Equally weighted		Value weighted	
Parameter	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Intercept	−0.038	−0.30	−0.002	−0.02	−0.512***	−3.30	−0.476**	−2.53
MRKT-rf	−0.061**	−2.50	−0.019	−0.85	0.056	1.53	−0.013	−0.27
SMB	−0.068	−1.14	0.134**	2.05	−0.281***	−3.36	−0.374***	−3.83
HML	−0.016	−0.32	0.051	0.91	0.140*	1.87	0.183**	2.41
MOM	0.080	1.23	−0.029	−0.86	0.042	0.86	−0.009	−0.14
N	156		154		156		156	
R-squared (%)	17.44		4.29		11.98		14.18	
Original DW			3.063					
Transf. DW			2.327					

Results from the Carhart model, decile portfolios

	ESGC score				ESG score			
	Equally weighted		Value weighted		Equally weighted		Value weighted	
Parameter	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Intercept	−0.349	−1.39	0.101	0.68	−0.366***	−2.71	−0.397**	−2.28
MRKT-rf	−0.089*	−1.79	−0.002	−0.05	0.097***	2.67	0.038	0.89
SMB	−0.137	−1.16	−0.124	−1.19	−0.376***	−4.04	−0.449***	−3.91
HML	0.012	0.13	0.260***	2.90	0.076	0.86	−0.002	−0.02
MOM	0.219	1.48	−0.069	−1.02	−0.021	−0.36	−0.021	−0.34
N	156		154		155		155	
R-squared (%)	21.68		13.08		17.07		10.33	
Original DW			3.306		2.701		2.575	
Transf. DW			2.332		1.937		1.988	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table provides the results from the Carhart model. The dependent variables are the excess return of four yearly rebalancing long-short portolios that mimics long good ESG and short bad ESG, zero investment. The first table shows the portfolios picked on top/bottom quartiles ESG performance, and the second table shows the portfolios picked on top/bottom deciles. On the left side of the tables, the scores are based on ESGC, and on the right side, they are based on ESG. Finally, the analysis covers both equally and value weighted portfolios. The variable MRKT-rf is the value weighted market return less the risk free rate, the SMB factor captures the portfolios exposure to small cap stocks and the HML factor captures the portfolios exposure to high book-to-market value firms. The MOM factor captures the portfolios exposure to previous winners or losers, i.e. the momentum, the coefficient picks up the difference in exposure between the good and the bad portfolio. Finally, the intercept captures the difference in abnormal return of the portfolios. We estimated the model with monthly data from 2005-2017.

Fama-French five-factor model

Table 17

Results from the Fama-French five-factor model, quartile portfolios

	ESGC score				ESG score			
	Equally weighted		Value weighted		Equally weighted		Value weighted	
Parameter	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Intercept	−0.050	−0.45	−0.016	−0.15	−0.483***	−3.15	−0.468**	−2.46
MRKT-rf	−0.069**	−2.37	−0.015	−0.63	0.057	1.58	−0.008	−0.18
SMB	−0.037	−0.65	0.128*	1.93	−0.267***	−2.90	−0.373***	−3.55
HML	0.004	0.04	0.063	0.62	0.073	0.78	0.149	1.25
RMW	0.176	1.38	−0.019	−0.14	−0.025	−0.19	−0.056	−0.32
CMA	0.055	0.54	−0.006	−0.05	0.106	0.71	0.051	0.27
N	156		154		156		156	
R-squared (%)	15.48		3.8		11.96		14.31	
Original DW			3.004					
Transf. DW			2.333					

Results from the Fama-French five-factor model, decile portfolios

	ESGC score				ESG score			
	Equally weighted		Value weighted		Equally weighted		Value weighted	
Parameter	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Intercept	−0.301	−1.62	0.173	1.09	−0.439***	−3.38	−0.480**	−2.51
MRKT-rf	−0.112*	−1.77	−0.001	−0.03	0.096**	2.52	0.035	0.80
SMB	−0.077	−0.69	−0.182*	−1.72	−0.364***	−3.70	−0.436***	−3.60
HML	−0.031	−0.14	0.198	1.34	0.180	1.56	0.127	0.74
RMW	0.283	1.23	−0.272	−1.39	0.156	0.94	0.192	0.79
CMA	0.122	0.56	−0.153	−0.79	−0.030	−0.24	−0.056	−0.33
N	156		154		155		155	
R-squared (%)	14.92		13.99		17.45		10.79	
Original DW			3.299		2.682		2.551	
Transf. DW			2.332		1.929		1.984	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table provides the results from the Fama-French five-factor model. The dependent variables are the excess return of four yearly rebalancing long-short portolios that mimics long good ESG and short bad ESG, zero investment. The first table shows the portfolios picked on top/bottom quartiles ESG performance, and the second table shows the portfolios picked on top/bottom deciles. On the left side of the table, the scores are based on ESGC, and on the right side, they are based on ESG. Finally, the analysis covers both equally and value weighted portfolios. The variable MRKT-rf is the value weighted market return less the risk free rate, the SMB factor captures the portfolios exposure to small cap stocks, the HML factor captures the portfolios exposure to high book-to-market value firms. The RMW factor captures the portfolios' exposure to firms with robust profitability, the coefficient picks up the difference in exposure between the good and the bad portfolio. The CMA factor captures the portfolios' exposure to firms with a conservative investment strategy, the coefficient picks up the difference in exposure between the good and the bad portfolio. Finally, the intercept captures the difference in abnormal return of the portfolios. We estimated the model with monthly data from 2005-2017.

Fama-French five-factor model with momentum

Table 18

Results from the Fama-French five-factor model with momentum, quartile portfolios

	ESGC score				ESG score			
	Equally weighted		Value weighted		Equally weighted		Value weighted	
Parameter	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Intercept	−0.096	−0.76	0.005	0.05	−0.502***	−3.06	−0.459**	−2.26
MRKT-rf	−0.063***	−2.64	−0.017	−0.73	0.060*	1.66	−0.009	−0.20
SMB	−0.054	−0.90	0.136**	2.05	−0.274***	−2.96	−0.370***	−3.50
HML	0.068	0.95	0.032	0.29	0.099	0.97	0.136	1.13
RMW	0.156	1.16	−0.019	−0.14	−0.033	−0.26	−0.052	−0.29
CMA	−0.013	−0.17	0.028	0.23	0.078	0.50	0.065	0.34
MOM	0.079	1.15	−0.033	−0.86	0.033	0.65	−0.016	−0.26
N	156		154		156		156	
R-squared (%)	18.49		4.37		12.29		14.36	
Original DW			3.004					
Transf. DW			2.327					

Results from the Fama-French five-factor model with momentum, decile portfolios

	ESGC score				ESG score			
	Equally weighted		Value weighted		Equally weighted		Value weighted	
Parameter	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Intercept	−0.432*	−1.89	0.209	1.20	−0.427***	−2.94	−0.471**	−2.30
MRKT-rf	−0.094*	−1.94	−0.006	−0.15	0.095**	2.48	0.034	0.77
SMB	−0.124	−1.05	−0.169	−1.58	−0.360***	−3.59	−0.433***	−3.56
HML	0.149	1.17	0.144	0.92	0.162	1.31	0.113	0.61
RMW	0.226	0.94	−0.273	−1.42	0.157	0.95	0.194	0.79
CMA	−0.069	−0.49	−0.094	−0.46	−0.011	−0.08	−0.041	−0.22
MOM	0.224	1.46	−0.057	−0.77	−0.020	−0.32	−0.016	−0.24
N	156		154		155		155	
R-squared (%)	22.51		14.74		17.57		10.84	
Original DW			3.299		2.682		2.552	
Transf. DW			2.344		1.934		1.986	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table provides the results from the Fama-French five-factor model with momentum. The dependent variables are the excess return of four yearly rebalancing long-short portolios that mimics long good ESG and short bad ESG. The first table shows the portfolios picked on top/bottom quartiles ESG performance, and the second table shows the portfolios picked on top/bottom deciles. On the left side of the table, the scores are based on combined score, and on the right side, they are based on ESG score. Finally, the analysis covers both equally and value weighted portfolios. The variable MRKT-rf is the value weighted market return less the risk free rate, the SMB factor captures the portfolios exposure to small cap stocks, the HML factor captures the portfolios exposure to high market-to-book value firms, the RMW factor captures the portfolios exposure to firms with robust profitability, the CMA factor captures the portfolios exposure to firms with a conservative investment strategy, and the MOM factor captures the portfolios exposure to previous winners and losers. Finally, the intercept captures the difference in abnormal return of the portfolios. We estimated the model with monthly data from 2005-2017.

9.4 Other tables

Portfolio assessment

Table 19

Average turnover of stocks in portfolios (%)									
	Small				Large				
	Good		Bad		Good		Bad		
	ESGC	ESG	ESGC	ESG	ESGC	ESG	ESGC	ESG	Average diff (%)
One-year	57.87	36.97	60.39	32.90	44.36	26.94	48.46	23.54	77.66
Three-year	68.90	46.71	61.10	39.98	49.44	33.46	50.88	29.12	55.71

Number of unique stocks that enters the portfolios over the total time span									
	Small				Large				
	Good		Bad		Good		Bad		
	ESGC	ESG	ESGC	ESG	ESGC	ESG	ESGC	ESG	Average diff (%)
Unique stocks	189	131	193	127	292	234	326	217	42.82

Average score volatility (σ)	
ESG score	31.74
ESGC score	43.09
Diff (%)	35.77

This table illustrates the different behaviours of the portfolios based on ESG-rating and combined score. The purpose is to provide relevant information about the portfolios that can contribute to increased understanding of the results from the analysis. As comes clear from the table, the combined score is a lot more volatile

STOXX and Fama-French

Table 20

	Fama-French	STOXX 600 Europe
Austria	✔	✔
Belgium	✔	✔
Switzerland	✔	✔
Germany	✔	✔
Denmark	✔	✔
Spain	✔	✔
Finland	✔	✔
France	✔	✔
Great Britain	✔	✔
Greece	✔	
Ireland	✔	✔
Italy	✔	✔
Netherlands	✔	✔
Norway	✔	✔
Portugal	✔	✔
Sweden	✔	✔
Poland		✔
Luxembourg		✔

This table presents the markets that are included in the construction of the Fama-French risk factors as well as the STOXX 600 Europe index. The small difference betseen the two confirms to a large extent that the European Fama-French factors are well suited for the analysis we are running.