LUNDS UNIVERSITET

Environmental, social and governance composite ratings and stock performance

Jesper Johansson

Tutors: Hossen Asgharian & Zahra Hashemzadeh

Department of Economics, Lund University, NEKH01, Bachelor's thesis

May 2019

Abstract

This paper studies the effect of non-financial performance on the stock performance of companies. By using ESG ratings from the Thomson Reuters Eikon DataStream we were able to sort the companies into portfolios of high and low ESG performance. The study finds that low scoring ESG companies significantly outperformed their high scoring counterpart for the period 2010-2015 but that the difference can partially be attributed to portfolio composition.

The study also studied the effect across economy sectors and the results were inconclusive for the period prior to the financial crisis but that the significant outperformance mentioned above drove the industries to all display negative results after the financial crisis. The study also finds that changes in ESG ratings have no predictive power and provide no new information. This would be in line with the efficient market hypothesis that ESG, as a public information do not provide any information that is not already priced.

Keyword: ESG, S&P500, ESG change , portfolio approach

Acknowledgements: I would like to express my deepest appreciation and gratitude to my supervisors Hossein Asgharian and Zahra Hashemzadeh for their valuable comments and suggestions during the writing process of this paper. I would also like to express my appreciation for Lund School of Economics for the privilege of attending their excellent educational program.

Lund May 2019

Jesper Johansson

Table of contents

1 Introduction

Over the last couple of years, the interest in socially responsible investing has skyrocketed with sustainable and impact assets expanded to account for $12 trillion in 2018 an increase of 38% from the $8,7 trillion in 2016. This is driven by asset managers and institutional investors incorporating environmental, social and governance (ESG) performance in their investing strategy. ESG rated assets has expanded by 44% from $8,1 trillion in 2016 to $11,6 trillion in 2018. The top three priorities for asset managers are climate change/carbon emissions, tobacco and minimizing conflict risk. The growing interest in socially responsible investing triggered the development of ESG ratings as a tool for screening companies' non-financial performance (US SIF, 2018).

ESG ratings as an objective measure of a company's non-financial performance is a relatively new phenomenon and this begs the question if non-financial performance have a significant impact on financial performance. This study aims to expand the knowledge on the effect of ESG on financial performance. The study uses a portfolio approach to sort and rank the stocks listed on the Standard & Poor's 500 index (S&P500) into portfolios based on their Thomson Reuters ESG ratings. A difference portfolio is constructed to measure the difference between high and low ESG stocks. The monthly average raw returns and the Carhart four-factor model (CH4) risk-adjusted returns are measured and compared for each portfolio. The effect on the financial performance is tested for holding periods of one and five years, for industry, ESG combined rating and for the annual change in ESG. Previous research shows that the non-financial performance can have three different effects on returns: The stocks that perform well on ESG outperform conventional stocks financially (Kempf & Osthoff, 2007), stocks that perform well on ESG underperform conventional stocks financially (Mollet & Ziegler, 2014) or that there is no benefit or cost to investing in companies that perform well on ESG (Limkriangkrai, Koh & Durand, 2017).

The study finds a significant negative effect of ESG on portfolio returns with negative and significant alphas for the difference portfolio even after risk-adjustment. The effect can partially be attributed to differences in style for the high and low ESG portfolios, but the significant alphas remain. This style difference and the significant underperformance of high ESG is somewhat mitigated but not removed entirely when using the ESG combined rating. The study finds no effect that can be attributed to the annual change in ESG ratings which suggests that ESG provides no new information. The negative effect of ESG on financial

performance is consistent for different economy sectors as well. The longer holding period portfolios suggest that after the financial crisis there is a 5-year period of consecutive outperformance for the low ESG portfolios but that this period ended in 2015. These results indicate that high ESG was outperformed by low ESG in the past but that the difference has disappeared.

This paper contributes to the topic by testing the effect of ESG ratings on a company level with a diverse selection of stocks. The paper also contributes by testing the effect of annual changes in ESG ratings. The diversified selection of stocks allows us to make inferences about the effect of ESG in a broad sense. The use of a composite ESG rating limits the study from testing the individual effect of the three pillars E, S and G. The use of the S&P500 limits the study to large cap stocks in the US.

The outline of this paper is as follows: section 2 discusses the theory and previous research. Section 3 discusses the structure of the data. Section 4 discusses the methods. Section 5 discusses the results of the data analysis. In section 6 conclusions are made and ideas for further research is discussed.

2 Literary review and theory

2.1 The history of ESG and ESG ratings

Socially responsible investing in the US has a long and illustrious history with roots tracing back to the famous sermon *the use of money* by the evangelist, cleric and co-founder of the Methodist movement John Wesley (Wesley Center Online, n.d). He provides three principles related to money management, gain as much as possible, save as much as possible and give back as much as possible without doing harm to others. The modern conception of socially responsible investing in the US gained momentum in the 60s with the civil and women's rights movements and the opposition towards the Vietnam war. Socially responsible investors aligned their capital allocation with their political and social beliefs with the hope that it would induce change (Donovan, 2019).

The idea of ESG investing was popularized with the publication of *Who Cares Wins* (UN Global Compact, 2005). The report asserts that better investment principles results in a better society. The publication led to the launch of the six principles of responsible investment (UNPRI, 2006). The six principals aim to develop a sustainable global financial system with the incorporation of ESG issues into investing.

ESG investing builds on the ethical and moral foundation of SRI investing and considers the non-financial performance as financially relevant. The ESG ratings are a set of standards that are used to measure a company's performance with regards to environmental, social and governance issues. The goal is to bridge the asymmetric information gap between companies and their share- and stakeholders. The rapid development of ESG can be seen by the growing number of ethical indices and companies that try to provide their own respective method of capturing ESG related performance and concerns. Sustainability reports over 600 different ESG ratings in 2019 globally, a five-fold growth since they published their first report in 2010 (SustainAbility, 2019).

2.2 Hypotheses: ESG effect on financial performance

The study tests the three hypotheses presented by Hamilton, Jo & Statman (1993) & Statman & Glushkov (2009) regarding the effect of non-financial performance on financial performance.

"Doing good but not well". The first hypothesis argues that companies that perform well according to ESG perform worse financially. This could be due to companies underestimating the cost or overestimating the benefit of ESG investing (Statman, 2006). This hypothesis echoes the thoughts of Friedman (1970) that a company that focuses on for instance ESG issues incur costs on society through its lack of focus and thus causing it to be less efficient. Friedman also argues that a company that maximizes shareholder value is the most socially responsible.

An example of a cost attributed to ESG would be Barnea & Rubin (2010) who show that ESG investing can provide an incentive for managers to promote ESG investment for personal gain where the cost exceeds the benefit at the expense of shareholders, thereby incurring an agency cost to ESG. Another explanation could be that investors tilting towards high performing ESG assets affect the stock price and drive up the valuation of the companies thereby lowering the returns (Statman, 2006).

"Doing good while doing well". The second hypothesis argues that investors who tilt towards companies that perform well on ESG concerns also receives a higher return than investors in conventional companies. Statman (2006) argues that this would be true if investors consistently underestimate the benefit or overestimate the cost of ESG investing, for instance Galema, Plantinga & Scholtens (2008) show that a high SRI score in employee relations deliver a positive and significant effect on company returns. Statman (2006) also argues that the high performing companies with regards to ESG could be undervalued as a result of not being favoured against their conventional counterpart.

"No effect". The third hypothesis argues that there should be no risk-adjusted difference in returns between companies that deliver a good or bad ESG performance. The efficient market hypothesis (EMH) argues that if a market is semi-efficient new public information like ESG ratings should be immediately incorporated in the price of the asset (Fama, 1970). If ESG ratings are shown to have no effect this would be consistent with the EMH and indicate that ESG ratings do not provide new information that is priced (Statman & Glushkov, 2009). This implies that the expected returns of a socially responsible company should be equal to its cost of capital because investors do not favour socially responsible companies (Statman, 2006).

Elzahar et al. (2015) show that the quality of financial and non-financial disclosure quality has a significant effect on the cost of capital for the former but not the latter. This could indicate that the non-financial disclosures are not priced. This hypothesis means that there should be no benefit or cost to investing according to ESG.

2.3 Previous research

Previous research on the effect of non-financial performance on financial performance is inconclusive. The subject is difficult since ESG as an objective measure of non-financial performance is a relatively new measure. Another problem with the ratings is that unlike financial key performance indicators (KPI), non-financial KPIs are unstandardized and can differ from one ratings company to another. This was shown by Elzahar et al. (2015) who also report that the unstandardized non-financial KPIs are of lesser quality than the financial KPIs. The study also shows that out the ten most used KPIs in the valuation process only two are non-financial. This indicates that the non-financial KPIs while valued are not yet as important as the financial KPIs. Coram, Mock and Monroe (2011) Show a similar result when studying the verbalized thought process of a small sample of Australian equity analysts. The study shows that non-financial KPIs are primarily given attention when a company is doing well while financial KPIs dominate attention during duress. The researchers note that non-financial KPIs are used primarily to increase confidence in an analysis. These results would indicate that ESG ratings are important but that analysts are not entirely certain how to use them and that they are of lesser importance then financial KPIs.

Another problem is that views among managers with regards to the importance of ESG investing differs across country borders (Duuren, Planting & Scholens, 2016). Ioannou & Serafeim (2017) echoes the result of differing views by showing that the propensity to disclose vary with legislation and the likelihood for repercussions in the case of non-compliance. Another problem with measuring the effect of ESG on financial performance derives from the structure of the ratings. Previous research shows that the composite rating and the three pillars E, S and G individually have different effect on the financial performance (See for example Limkriangkrai, Koh and Durand, 2017, Buallay, 2019, Scholtens & Zhou, 2008)

Kempf & Osthoff (2007) uses SRI data from KLD research & analytics to test the effect of SRI on financial performance. Several different trading strategies are used based on the SRI sorting. For instance, a long position in high SRI companies and short position in low SRI companies, commonly referred to as a high-minus-low (HML) portfolio. The study also uses a negative exclusionary approach that shuns stocks that are considered unethical (Ex. Tobacco, alcohol & military). The study find that high SRI portfolios display a positive and significant abnormal return but that a strategy of excluding stocks through a negative screening process yields a negative abnormal return.

Statman & Glushkov (2009) show a similar result with a trading strategy that tilt towards socially responsible stocks. The socially responsible portfolio significantly outperformed the conventional investment strategy. The study also reports the same results for a strategy that shun stocks based on a negative screening process, a shunning strategy together with a tilt significantly underperformed the conventional strategy. Statman (2006) show that SRI indices outperformed the S&P500 for the period prior to the dot-com bubble but that the indices reversed after the bubble collapsed. Buallay (2019) show that the composite ESG rating has a significant positive effect on the financial performance of European banks measured as Tobin's Q, return on equity (ROE) and return on assets (ROA).

Several other studies find a negative effect of ESG ratings and other ESG related measures on financial performance. Mollet & Ziegler (2014) compared the performances of companies regarded as SRI leaders with ordinary companies in the US and Europe for the period 1998-2009. The study found after risk-adjustment that ordinary companies outperformed the SRI leaders in both regions but that it was only significant for Europe. Belghitar, Clark & Desmukh, (2014) found no significant difference between investing in conventional and SRI indices. The study also questioned the quadratic utility of the modern portfolio theory assumed in previous studies and developed a model that incorporates the higher moments, skewness and kurtosis. The model showed that when higher moments are incorporated the SRI indices performs significantly worse than their conventional counterpart and pays a premium due to lower skewness and higher kurtosis.

Hong & Kacperczyk (2009) report a "sin"-premium associated with stocks that are shunned by norm constrained institutional investors and other investors with fiduciary duties. The neglect by these investors led to the "sin"-stocks being undervalued with higher book-to-market ratios which results in a significant outperformance compared to conventional stocks. Das et al. (2018) show that high ESG rating among socially responsible mutual funds (SRMF) display significantly better risk-adjusted returns than their low ESG rating counterparts during the financial crisis. However, during the periods prior to and after the financial crisis the low and mid ESG rating SRMF display significantly better risk-adjusted returns than their high rating counterparts.

El Ghoul & Karoui (2017) report similar results when comparing high corporate social responsibility (CSR) equity funds with conventional funds. High rated CSR equity funds display a negative and significant alpha when compared to lower CSR rated equity funds, however the high CSR equity funds also display a stronger performance resistance. This result

is consistent with the discriminatory taste argument in Manescu (2011) that socially responsible investors derive non-financial utility from investing in businesses regarded as ethical.

Previous studies have also found that non-financial performance have no effect on financial performance. Humphrey, Lee & Shen, (2012) tests the effect of ESG ratings on the returns of stocks listed on the UK stock market. The study finds no differences in returns or risk for companies based on their difference in ESG, nor do they find any differences when controlling for industries.

Limkriangkrai, Koh & Durand (2017) use a portfolio approach to measure the effect of ESG on financial performance on the Australian stock market. The study found no significant effect for the composite ESG ratings even before risk-adjustment. Landi & Sciarelli (2019) show similar results studying the Italian stock market and reports no significant benefit or cost to companies that perform well on ESG. Unlike Das et al. (2018) an older study of SRMF finds no significant over/underperformance (Hamilton, Jo & Statman, 1993)

3 Data and empirical material

3.1 Stock screening

To measure the effect of the ESG ratings on the financial performance of companies, the study uses the monthly stock prices of the companies listed on the S&P500. The stocks prices are obtained from the Thomson Reuters Eikon DataStream for the period 2003-2019. The log-returns are calculated using the following formula.

$$r_{i,t} = \ln(P_{i,t}) - \ln(P_{i,t-1})$$

The log-returns are used, as oppose to the simple returns due to the desirable properties of the log-returns. If prices are assumed to be log-normally distributed, then log-returns are normally distributed which is a good property for statistical inference. The stock price as a measure of return have been used in previous studies (see for example Kempf & Osthoff 2007, Landi & Sciarelli, 2018). An alternative approach would be to use accounting-based performance measures like ROE or ROA (see for example Buallay, 2019).

3.2 S&P500

The study uses the stocks of the companies listed on the S&P500. Prior to analysis, stocks that lack the required monthly price data are removed. This step reduced the sample from 505 stocks originally to 501 stocks. The S&P500 was chosen for the study based on its diversified nature,

the index is by design constructed to be representative of the American industry. This is advantageous because the diversified nature of the index covers a large spectrum of industries in the investable universe. The wide variety of stocks in different industries also means the study will benefit from having a large spread in non-financial performance. An alternative approach would be to use the stocks of one of the indices constructed using the best ESG related performers. An example would be for instance to use stocks listed on the *MSCI KLD 400 social index* which is constructed to encompass and expose companies regarded as top ESG performers. This has been done in previous studies (see for example Belghitar, Clark & Desmukh, 2014, Statman, 2006).

The S&P500 is regarded as large enough to provide some insights about the effect of non-financial performance on financial performance. Use of the S&P500 is disadvantageous in that it only encompasses US large-cap stocks. This limits the study to making inferences about the effect on US large cap mainly. To investigate if there is a difference between industries in the effect of ESG ratings, the stocks are sorted according to the Thomson Reuters Business classification (TRBC) economic sectors. Since a company's ESG rating is benchmarked relative to the TRBC industry group a comparison based on the subdivision economy group will not be as robust as a comparison on the industry level. However, a subdivision based on industry group would result in many small portfolios which makes inference impossible using only the S&P500. We argue that a subdivision based on the economy group is fair because the companies within each group has more in common with each other than other groups. From an investor's perspective a high ESG rating irrespective of it belonging to one economy group or another is still a sign of a company performing well with regards to ESG. We argue that if capital allocation based on ESG ratings is used within the screening process to induce positive changes within the ESG criteria, the best performing companies should benefit irrespective of industry.

The S&P500 only represent the US industry, this allows the study to make inferences about the US market. These results however might not be applicable to other countries with different corporate structures, markets and views on the non-financial performance. As shown by Ioannou & Serafeim (2017) the propensity to disclose ESG performance and the subsequent quality of disclosure varies across countries. This difference can be explained by legislatively mandated disclosures or by differing views on the necessity of ESG disclosure. Duuren, Plantinga & Scholtens, (2016) for instance found that the views on ESG differs for US and European managers.

3.3 ESG Data

The purpose of this study is to investigate the effect of a company's non-financial performance on their financial performance. The study uses a portfolio approach which requires loadings on a sorting variable. The assets are ranked and sorted into portfolios based on their loading on the sorting variable. Previous studies on the effect of non-financial performance ratings on financial performance use a wide variety of different ESG/SRI ratings. Many use the MSCI KLD ESG ratings (see for example Galema, Plantinga & Scholtens, 2008, Kempf & Osthoff 2007, Scholtens & Zhou 2008, Statman & Glushkov 2009). Another study uses the ratings provided by Morningstar (Das et al 2018), Sustainable asset management (Humphrey, Lee & Shen 2012) and another study use the ratings from regional providers Regman ESG rating (Limkriangkrai, Koh & Durand 2017).

The large number of different ESG ratings causes a problem when comparing different studies. The ratings companies use different ideas, methods and datapoints to calculate a company's non-financial performance. For instance, Thomson Reuters ESG score use 400+ datapoints to calculate a company's relative rating within the ten themes (Refinitiv, 2019) while for instance MSCI use 1000+ datapoints to calculate a company's performance in 37 key issues (MSCI, 2018). Another issue is that non-financial key performance indicators (KPI) are not standardized to the degree of for instance financial KPIs and have been shown to have lower quality (Elzahar et al, 2015). The unstandardized KPIs causes an inherent difficulty for companies when compiling disclosures and even more so for the ratings companies when they evaluate said disclosures.

This study uses for the sorting variable and measure of non-financial performance the Thomson Reuters (formerly Asset4) ESG and ESG combined ratings. The ratings for the period 2003-2018 are obtained from the Thomson Reuters Eikon DataStream. Stocks that lack the necessary ESG ratings are not removed. Instead, they are handled within the confines of the portfolio approach. The Thomson Reuters ESG ratings are composite ratings of a company's relative performance across ten main themes (resource use, emissions, innovations, workforce, human rights, community, management, shareholders, CSR strategy). The ten themes make up the three pillar ratings E, S and G. The composite ESG ratings show the company's percentile rank benchmarked against their respective industry group. The ratings are based on a company's public information and self-reported ESG related disclosure. The ESG combined rating is a combination of the company's ESG and ESG controversy rating where controversy is a rating

of ESG related concerns. If a significant and widely covered negative ESG related event occurred during the period, the controversy rating penalises the ESG combined rating.

A concern with the Thomson Reuters ESG ratings is that the company-based ratings are relative to their respective Thomson Reuters business classification industry group benchmark. This is a strength and a weakness with the ratings. An investment strategy screening for high ESG ratings results in an industry diversified portfolio containing the best performing companies relative to their industry group. It is a weakness because the relative nature of the ratings might be confusing and even illusive to an investor. A portfolio with a high rating might contain stocks that are not usually associated with a socially responsible, ethical or an environmental investing approach.

Another problem with using the Thomson Reuters ESG composite ratings is the lack of information on the three pillar ratings. The Thomson Reuters Eikon DataStream only provides the composite ESG, ESG controversies and the ESG combined ratings. The limitation of using only composite ratings inhibits the study from testing the effect of the pillar ratings individually on the financial performance. Previous studies have shown that individual pillar ratings have different effect on financial performance (see for exampla Buallay, 2018, Limkrinangkrai, Koh & Durand, 2017, & Galema, Plantinga & Scholtens, 2008).

3.4 Data for common risk factors

To represent the risk-free rate, market risk, size, value and momentum factors in the CAPM and multifactor regression the study uses the data provided by Kenneth. R French's data library[1]. The data is necessary to calculate the excess and risk-adjusted returns of the portfolios. The data uses for its risk-free rate the one-month treasury bill rate. The market risk factor is defined as the excess return of a value weighted portfolio of all stocks on the NYSE, AMEX and the NASDAQ.

The size risk factor is defined as

$$SMB = {}^{1}/_{3}\,(Small\ Value + Small\ Neutral + Small\ Growth) - {}^{1}/_{3}\,(Big\ Value + Big\ Neutral + Big\ Growth)$$

The value risk factor is defined as

$$HML = {}^{1}/_{2}\,(Small\ Value + Big\ Value) - {}^{1}/_{2}\,(Small\ Growth + Big\ Growth)$$

[1] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/index.html

The momentum risk factor is defined as

$$MOM = {}^{1}/_{2}\,(Small\ High + Big\ High) - {}^{1}/_{2}\,(Small\ Low + Big\ Low)$$

The common risk factors are constructed by using a self-financing portfolio that tries to mimic the movements of the factors. The portfolio approach is flawed in that it only uncovers if there is an effect of the non-financial performance on the financial performance. The cause of the difference is not explained solely by the portfolio approach. Using the common risk factors in a regression approach results in estimates of the covariation of the portfolios with the common risk factors, loadings. The loadings on the common risk factors can provide some information about the portfolio composition and if there is a difference between the portfolios. The estimates of the market risk, and common risk factors have been estimated by using the American stock market which makes it a suitable match with the sample data retrieved from the S&P500.

3.5 Descriptive results

The growing importance of ESG for investors and the growing availability and comparability of data on non-financial performance can be deduced from the ratio of companies on the S&P 500 that is assigned an ESG rating. The number of companies that perform the necessary ESG disclosures and therefore are assigned an ESG ratings has been steadily increasing over time as can be seen in figure A.1 in the appendix. In 2003 less than 50% of all companies on the S&P500 had an ESG rating while in 2016 all companies were assigned an ESG rating.

Another sign of the growing importance of ESG and companies seemingly taking ESG related issued seriously is that the ESG ratings have been trending upwards during the time period. Figure A.2 shows the median, the 80^{th} and 20^{th} percentile ESG ratings. As can be seen there is a small upward trend among companies to become better at complying with ESG related issues. This could be due to the companies becoming better at filing the necessary ESG disclosures or that they walk-the-talk and perform the necessary steps to perform better according to ESG standards.

4 Method

4.1 Portfolio approach

The portfolio approach is a factor mimicking approach that try to capture the effect of an underlying factor on return variation (Fama & French, 1993). The method has been used in previous studies to examine the effect of non-financial performance on financial performance (see for example Limkriangkrai, Koh & Durand, 2017). The method is advantageous because it is a non-parametric statistical method (Bali, Engle & Murray, 2016). This implies that is does not make any assumptions about the nature of the effect of the factor loadings on the return variation.

The assets are ranked based on the factor loadings and sorted into portfolios. A difference portfolio is created by taking a long and a short position in the high and low loading portfolios respectively. The difference portfolio is a zero-investment strategy that is assumed to be sensitive to the underlying factor that makes up the sorting variable. The difference portfolio has an expected return of zero if there is no effect of the factor loadings on the return variation. In the case of no effect the return variation of the long and short position will counteract each other. A significant difference portfolio indicate that the loadings of the sorting variable has some effect on return variation.

The sorting variable is defined as X_t. The sorting variable is used to sort the stocks into decile, quintile, tercile and 2-quantile portfolios depending on the dispersion in the sorting variable and the size of the sample. This study uses several different sorting variables to gain insight in the effect of non-financial performance on financial performance. To test the effect of non-financial performance on financial performance for the one- and five-year portfolios and the effect for different economy sectors the study use the companies ESG ratings. The study also uses the ESG combined rating to test if any effect remains with the rating that punishes companies in accordance with ESG controversy. To test the predictive capacity of ESG ratings the annual change in the ratings is used as the last sorting variable. A company's ESG rating is treated as the factor loading on the stock that represents how well the company performs on ESG related topics

The breakpoints are the distinctions in the dispersion of the sorting variable where the company's factor loading on the sorting variable decide which portfolio it belongs to. For decile portfolios the breakpoint is the values of the sorting variable at the 10^{th},…,90^{th} percentile. There

is a trade-off between the number of stocks in each portfolio and the number of portfolios. Too many breakpoints result in small portfolios with a lot of intrinsic risk and noise while too few breakpoints result in too few portfolios with a small difference in the dispersion of the ESG rating. Too few breakpoints make it difficult to infer anything about the effect of the sorting variable on the return variation. According to early research by (Evans & Archer, 1968) at least ten stocks are needed for the intrinsic risk of a portfolio to be minimized while (Statman, 1987) argues that the minimum number of stocks should be thirty. In the study we try to obtain portfolio sizes above thirty stocks when possible.

For the test of the effect of ESG ratings on financial performance ten portfolios was an appropriate number of portfolios. This means that nine breakpoints had to be calculated for each period t. We define breakpoint k at time t as $B_{k,t} = Pctl_{p,k}(\{X_t\})$ where $Pctl_p(Z)$ = is the pth percentile of the set Z and $\{X_t\}$ is the set of valid values of the sorting variable. Companies that lack an ESG rating was excluded from the calculation of the breakpoints. The companies are ranked based on the sorting variable at time t, X_t and sorted into the kth portfolio $P_{k,t}$ based on their percentile rank. The portfolios are equally weighted giving each stock the weight $1/n_t$ in each portfolio. The portfolios are held for 12 months until a new set of the sorting variable is released at which point the portfolios are rebalanced in accordance with the new values of the sorting variable.

To measure the performance of the portfolios $1 \dots n_p$, the equally weighted average of the log-returns for each month is calculated according to.

$$\bar{r} = \frac{\sum_{i \in P_{k,t}} W_{i,t} r_{i,t}}{\sum_{i \in P_{k,t}} W_{i,t}} \; for \; k \in \{1, \dots, n_p\}$$

We also define the monthly return of the difference portfolio as the difference between the average monthly return of the high and low scoring ESG portfolios: $\bar{r}_{Diff,t} = \bar{r}_{n_{p,t}} - \bar{r}_{1,t}$. The one year holding period portfolios are compiled into a time series of portfolios rebalanced annually in accordance with the updated values of the sorting variable. The time series mean of the portfolios and the difference portfolio is defined as:

$$\bar{r}_k = \frac{\sum_{t=1}^{T} \bar{r}_{k,t}}{T} \text{ and } \bar{r}_{Diff} = \frac{\sum_{t=1}^{T} \bar{r}_{Diff,t}}{T}$$

The significance of the time series means of the portfolios report if the portfolios have earned a significant monthly average return over the time period $t = 1, \dots, T$. The time series

mean of the difference portfolio is the basis of the test of the effect of the sorting variable on the return variation. The difference portfolio is a zero-investment strategy that is assumed to be sensitive to the underlying factors that determine the ESG ratings which means that:

$$H0:\ \bar{r}_{Diff,t} = 0$$

$$H1: \bar{r}_{Diff,t} \neq 0$$

If the null hypothesis is rejected there is significant evidence of a significant effect of the sorting variable on the return variation. A significantly different result can be confirmed by plotting the time series mean of the portfolios along across the portfolio universe from high to low. If a monotonic trend is discernible in the time series as the plot progresses across portfolios there is evidence of an effect of the sorting variable on the return variation, if there is no trend the results could be spurious.

4.2 Capital Asset Pricing Model (CAPM)

The CAPM was developed independently by multiple economists (Sharpe 1964, Mossin 1966) building on the work of (Markowitz, 1952). CAPM tries to explain the expected return of an asset with the asset's sensitivity to the market portfolio. According to the model an asset is subject to two types of risk the systematic risk represented by the β_i and the idiosyncratic risk of that particular asset. The systematic risk is economy wide and cannot be diversified away while the idiosyncratic risk can be diversified away by adding more assets to the portfolio. A rational investor should hold a large enough portfolio as to eliminate idiosyncratic risk through diversification. The expected return of a portfolio should only be dependent on the amount of systematic risk in the portfolio.

$$r_{i,t} - r_{f,t} = a_{i,t} + \beta_i\left(r_{m,t} - r_{f,t}\right) + \varepsilon_{i,t}$$

Where $r_{i,t}$ is the return of asset i for time t, $r_{f,t}$ is the risk-free rate of return, $a_{i,t}$ is Jensen's alpha, β_i is the sensitivity of asset i to the market portfolio, $r_{m,t}$ is the return of the market portfolio and $\varepsilon_{i,t}$ is the error term. If the CAPM holds then all risk-adjusted reward of the asset should be explained by the asset β_i. This results in the portfolio alpha being insignificantly different from zero. If alpha is not zero, then the portfolio significantly over/underperforms what the CAPM predicts and the CAPM cannot hold. The CAPM is used in conjunction with the FF3 and CH4-factors to try to explain the returns of the decile portfolios and HML-portfolio. Just like (see for example Humprey, Lee & Shen 2012, Kempf & Osthoff

2007) the alphas of the portfolios are of interest since they represent the unexplained excess return that is associated with a certain exposure to the ESG factor.

4.3 Fama-French 3-factor model

The FF3 model is an extension of the CAPM that adds the two common risk factors size and value. (Fama & French, 1992).

$$r_{i,t} - r_{f,t} = a_{i,t} + \beta_{1,i}(r_{m,t} - r_{f,t}) + \beta_{2,i}SMB + \beta_{3,i}HML + \varepsilon_{i,t}$$

Where SMB and HML is the size and value risk factor respectively. The factors were constructed when it was observed that two types of stocks tended to out(under)perform their counterparts and the market, small(large)-cap stocks and values(growth) stocks. The factors are constructed to act as proxies for the two risk factors and are used in conjunction with the market risk factor to explain the expected returns of the portfolios.

4.4 Carhart four-factor model

The CH4 model is an extension of the CAPM and FF3 model (Carhart 1997). The model adds a fourth factor momentum to the FF3 extension of the CAPM. The factor was added by Carhart after the observation that stocks that have risen(fallen) in the past tended to continue trending up(down)wards (Jegadeesh & Titman 1993).

$$r_{i,t} - r_{f,t} = a_{i,t} + \beta_{1,i}(r_{m,t} - r_{f,t}) + \beta_{2,i}SMB + \beta_{3,i}HML + \beta_{4,i}MOM + \varepsilon_{i,t}$$

Where MOM is the factor representing the momentum risk factor. The CH4 model is used in conjunction with the CAPM and FF3 model to estimate the alphas of the portfolios. The model is also used to estimate the loadings of the common risk factors on the portfolios.

4.5 The common risk factor regressions

To obtain the performance of the portfolios adjusted for the four common risk factors specified in the CAPM, FF3 and CH4 models, a staggered OLS regression approach is used with the monthly excess returns of the portfolios as the dependent and the common risk factors as the independent variables added in three successive regressions. A portfolios abnormal return is measured as a significant alpha, Jensen's alpha. In addition to testing the risk-adjusted performance the regression displays the loadings of the common risk factors on the portfolios. The loadings can give some information about differences in portfolio composition across the portfolios.

The excess returns were used for the decile portfolios sorted on the sorting variable since it is assumed that they were purchased with the proceedings of borrowing at the risk-free rate. For the difference portfolio the monthly difference return was used since it is assumed that the high factor loading portfolio was purchased with the proceedings of short selling the low factor loading portfolio. For robustness the significance test of the regression coefficients uses the Newey-West robust standard errors to correct for autocorrelation and heteroscedasticity.

4.6 Performance measures

Jensen's alpha

Jensen's alpha is a measure of the abnormal return of an asset or a portfolio (Jensen, 1967). The measure compares the expected return of a portfolio with observed return by measuring the residual return commonly referred to as the abnormal return, $a_{i,t}$.

$$r_{i,t} = a_{i,t} + r_f + \beta_{1,i}(r_{m,t} - r_f) + \varepsilon_{i,t}$$

Where $a_{i,t}$ is Jensen's alpha. The formula above represents Jensen's alpha in a CAPM setting but it is generalisable and common to add for instance the FF3, CH4 factors or more. The CAPM asserts that the excess return of an asset or a portfolio should be equivalent to the excess return of the market portfolio multiplied by the assets systematic risk represented by the beta. If CAPM holds, then the abnormal return represented by $a_{i,t}$ should be equal to zero. If $a_{i,t} \neq 0$ then the portfolios are displays abnormal returns relative to the risk factors included in the regression.

Sharpe-Ratio

The Sharpe-Ratio is a commonly used ratio when comparing the risk-adjusted returns of single assets or portfolios of assets. The ratio was introduced by Sharpe to measure the reward-risk ratio (Sharpe, 1966)

$$SR_i = \frac{(r_{i,t} - r_f)}{\sigma_i}$$

Where σ_i is the standard deviation of asset i. The ratio measures the reward to volatility ratio of an asset.

5 Results

5.1 ESG ratings effect on returns 2003-2018

Table 5.1.1

Descriptive statistics for the decile portfolios for the period 2003-2018

	High-10	9	8	7	6	5	4	3	2	Low-1	Diff
Average raw return	0,78%	0,69%	0,76%	0,84%	0,98%	0,74%	0,79%	1,09%	0,92%	1,26%	-0,47%
Standard deviation	0,046	0,051	0,049	0,053	0,053	0,051	0,051	0,049	0,052	0,051	0,022
Skew	-1,79	-2,27	-2,22	-2,46	-2,53	-1,89	-2,09	-1,98	-2,90	-2,98	0,56
Kurtosis	8,76	12,76	12,97	14,47	14,26	10,23	11,15	10,40	20,70	21,27	4,96
Sharpe-ratio	0,15	0,12	0,14	0,14	0,17	0,13	0,14	0,20	0,16	0,23	-0,21

Notes:
N: of monthly observations: 180

The tables A.3 & A.4a-b in the appendix reports the number of stocks in each decile portfolio and their respective ESG ratings for each year. Table 5.1.1 reports some summary statistics about the decile portfolios sorted on ESG ratings and the difference portfolio for the period 2003-2018. The monthly average raw returns are higher for the portfolios with lower ESG ratings. A monotonic trend is observed when moving from the higher ESG portfolios to the lower ESG portfolios (see figure A.5), this indicates that the result is not spurious. The difference portfolio confirms this result with a negative monthly average raw return.

The standard deviation is congruent across the decile portfolios with no remarkable difference. The low ESG portfolio reports the highest Sharpe-ratio. All decile portfolios show a negative skewness and large kurtosis. Unlike Belghitar, Clark & Desmukh (2014) we find that higher ESG results in a higher skewness and a lower kurtosis compared with the lower ESG. The large values on skewness and kurtosis is shown to be caused by including the confounding effect of the financial crisis see table 5.2.1. As expected for a zero-investment strategy the difference portfolio displays lower standard deviation, skewness and kurtosis.

Table 5.1.2

Panel A: Monthly average portfolio raw returns and alphas 2003-2018

	High-10	9	8	7	6	5	4	3	2	Low-1	Diff
Average Return	0,78%	0,69%	0,76%	0,84%	0,98%	0,74%	0,79%	1,09%	0,92%	1,26%	-0,47%
	(2.26)**	(1.81)*	(2.10)**	(2.10)**	(2.49)**	(1.95)*	(2.10)**	(2.99)***	(2.37)**	(3.29)***	(2.87)***
CAPM Alpha	-0,19%	-0,34%	-0,25%	-0,26%	-0,11%	-0,30%	-0,25%	0,08%	-0,13%	0,26%	-0,44%
	(1.67)*	(2.02)**	(1.59)	(1.54)	(0.54)	(2.01)**	(1.44)	(0.52)	(0.61)	(1.17)	(2.52)**
FF3 Alpha	-0,19%	-0,33%	-0,25%	-0,25%	-0,10%	-0,31%	-0,24%	0,07%	-0,13%	0,24%	-0,43%
	(1.74)*	(2.07)**	(1.62)	(1.50)	(0.48)	(1.98)**	(1.37)	(0.47)	(0.61)	(1.11)	(2.60)**
Carhart 4 Alpha	-0,18%	-0,31%	-0,23%	-0,23%	-0,07%	-0,26%	-0,21%	0,09%	-0,11%	0,26%	-0,44%
	(1.59)	(1.85)**	(1.44)	(1.31)	(0.34)	(1.74)*	(1.16)	(0.57)	(0.50)	(1.21)	(2.61)***

Panel B: Loadings on the on the common risk factors using the Carhart four-factor model

Factors	High-10	9	8	7	6	5	4	3	2	Low-1	Diff
RM	1,021	1,070	1,053	1,114	1,100	1,057	1,019	1,029	1,065	1,036	-0,015
	(15.16)***	(10.20)***	(11.64)***	(10.67)***	(9.57)***	(12.54)***	(10.87)***	(11.8)***	(8.35)***	(7.61)***	(0.17)
SMB	-0,091	-0,118	-0,035	0,018	0,046	0,050	0,121	0,108	0,119	0,124	-0,215
	(1.41)	(1.15)	(0.50)	(0.22)	(0.54)	(0.63)	(1.48)	(1.38)	(1.32)	(1.36)	(2.78)***
HML	-0,046	0,020	-0,088	0,024	0,036	-0,093	0,082	-0,100	-0,088	-0,281	0,234
	(0.44)	(0.21)	(1.04)	(0.28)	(0.36)	(0.90)	(0.93)	(0.86)	(0.97)	(2.33)**	(3.21)***
MOM	-0,051	-0,093	-0,093	-0,085	-0,098	-0,132	-0,108	-0,069	-0,086	-0,088	0,037
	(0.70)	(1.52)	(1.31)	(1.44)	(1.70)*	(2.42)**	(1.90)*	(1.27)	(1.48)	(1.36)	(1.20)

Notes:
T-stat is reported in the brackets
*Asterix denotes * p<0,10, ** p<0,05, *** p< 0,01*
N. of monthly observations 180

Panel A of Table 5.1.2 reports the monthly average returns for the decile portfolios and difference portfolio with t-statistics. The returns are significantly different from zero for all decile portfolios and the difference portfolio. The negative and significant difference portfolio indicates that the portfolio containing low ESG rated stocks significantly outperformed the portfolio with high ESG rated stocks. In figure A.10 the monthly average for each year is plotted, these two portfolios make up the difference portfolio and the average returns seem to track each other for the period. The Low-1 portfolio consistently outperforms or performs at least as good as the High-10 portfolio. After the financial crisis it consistently outperforms the High-10 portfolio up until the year 2015. The period 2010-2015 could be what drives the average raw returns of the difference portfolio to be significant.

Panel A of table 5.1.2 also reports the CAPM, FF3 and CH4 alphas of the portfolios. The alphas are obtained by performing an OLS regression with the excess (raw) returns of the decile (difference) portfolios as the dependent variable and the common risk factors as the independent variables. A significant alpha indicates that the portfolio has some excess return that cannot be explained by the common risk factors. A significant positive (negative) alpha is usually a positive(negative) sign for an investment since it outperforms(underperforms) what is expected according to the CAPM and multifactor models. All decile portfolios except the portfolios 3 and Low-1 display a negative alpha during the time period 2003-2018. The alphas

are negative and significant for the two highest ESG portfolios and a middle portfolio. The difference portfolio displays negative and significant CAPM, FF3 and CH4 alphas. This further indicates that there is a significant difference between the High-10 and Low-1 portfolio that is not explained by the market risk or the common risk factors.

Panel B of table 5.1.2 reports the loading of the decile and difference portfolios on the CAPM, FF3 and CH4 factors. All decile portfolios display a positive market beta coefficient (RM) above but close to one. This indicates that the decile portfolios closely track the movements of the market portfolio factor which per definition has a coefficient of one. The difference portfolio has a negative and insignificant market beta coefficient. This is expected since the difference portfolio is a zero-investment strategy where the long and short portfolio position nullify each other.

The decile portfolios have insignificant loadings on the SMB factor but there is a monotonic trend across the portfolios. The high ESG portfolios have a negative loading on the SMB factor but as ESG ratings become lower the factor loading on the portfolios becomes positive. This indicates a difference in portfolio compositions, that the ESG ratings have a positive relationship with market capitalization. Since the S&P500 consists entirely of large cap stocks the increasing loadings on the SMB factor indicates that the lower ESG portfolios consist of stocks that act similarly to small cap stocks. This result is expected since previous research show that ESG is positively related to size (see for example Das et al. 2018, Hong & Kacperczyk, 2009 & Humphrey, Lee & Shen, 2012). The difference portfolio confirms the difference in loading on the SMB factor by being negative and significant. This is explained by the long and short positions in the High-10 and Low-1 portfolios respectively. If the high and low ESG portfolios are tilted towards large- and small cap respectively as is indicated by the factor loadings, the portfolio composition of the difference portfolio would increase the tilt with the long- and short composition. This result indicate that the negative and significant alphas can possibly be explained by a difference in exposure to the risk factor size.

All decile portfolios except the Low-1 portfolio have an insignificant loading on the HML factor. Unlike the SMB factor there is no clear monotonic trend for the loadings on the HML factor. The High-10 and Low-1 portfolios both displays negative factor loadings where the former is insignificant, and the latter is significant. This indicates that the portfolios composition of the Low-1 portfolio is significantly tilted towards growth stocks (low B/M) while the High-10 portfolio is more neutral but relative to the Low-1 it is more tilted towards value stocks (high B/M). Humphrey, Lee & Shen (2012) found similar tilts for ESG sorted

portfolios. The difference portfolio has a positive and significant loading on the HML factor. The short position in the negative and significant Low-1 portfolio outweighs the long negative and insignificant position in the High-10 portfolio. This results in the difference portfolio having a positive and significant tilt towards the HML factor. The difference portfolio has a significant exposure to the value premium. All portfolios except the difference portfolio display a negative loading on the momentum factor with three portfolios having a large enough tilt to be significant. The negative loadings indicate that the stocks in the decile portfolios exhibit poor past performances. The portfolios are composed of stocks associated with a contrarian investment strategy. A negative factor loading on the momentum factor indicates that the decile portfolios are missing out on the momentum premium observed by Jegadeesh & Titman, (1993). The difference portfolio displays a positive but insignificant tilt towards the momentum factor.

There is significant difference in returns when comparing high and low ESG rated companies. The returns show a negative effect of ESG on financial performance. The tests show that this could be due to differences in portfolio composition. ESG is negatively associated with the size factor while the difference portfolio shows that there is also a significant different in the loading on the value factor. The figure A.6 shows the monthly returns of the difference portfolio. The returns are stochastically dispersed around the zero with a level that seems to be slightly below zero for the period before the financial crisis and the period 2010-2015. This result confirms the previous result of a slight performance advantage for the Low-1 portfolio.

A more intuitive approach to illustrate the performance difference between the High-10 and Low-1 portfolios is the figure A.7. The figure plots the development of a $1 investment in the difference portfolio. Between the period 2003-2015 the value of the $1 investment is consistently downward trending with exceptions for a few upward jumps prior to and during the financial crisis, after 2015 the negative trend seems level out and disappear completely. The zero-investment strategy should if there is no difference between high and low ESG portfolios result in a negligible difference. The results above indicate that a portfolio tilted towards high ESG firms significantly underperforms a portfolio that is tilted towards low ESG. This effect however seems to have disappeared post 2015.

5.2 ESG ratings effect on returns prior to and after the financial crisis

To test if the negative effect of ESG ratings on financial performance is consistent when excluding the confounding effect of the sub-prime mortgage crisis of 2007-2009 the sample is split into two subsets: pre- and post-financial crisis.

Table 5.2.1

Descriptive statistics for the decile portfolios for the period 2003-mid 2007 (pre-financial crisis)											
	High-10	9	8	7	6	5	4	3	2	Low-1	Diff
Average raw return	1,22%	1,50%	1,28%	1,50%	1,90%	1,24%	1,53%	1,39%	1,43%	1,68%	-0,46%
Standard deviation	0,027	0,026	0,030	0,030	0,028	0,031	0,028	0,030	0,032	0,031	0,018
Skew	-0,28	-0,02	-0,05	-0,24	-0,31	-0,14	-0,27	-0,33	-0,12	0,06	-0,19
Kurtosis	-0,45	-0,25	-0,17	-0,47	-0,26	-0,10	0,00	-0,47	-0,46	-0,41	0,65
Sharpe-ratio	0,37	0,48	0,35	0,42	0,59	0,33	0,46	0,39	0,38	0,46	-0,25
Descriptive statistics for the decile portfolios for the period mid 2009-2018 (post-financial crisis)											
	High-10	9	8	7	6	5	4	3	2	Low-1	Diff
Average raw return	1,18%	0,93%	1,15%	1,16%	1,33%	1,15%	1,23%	1,53%	1,50%	1,66%	-0,48%
Standard deviation	0,040	0,041	0,039	0,042	0,043	0,041	0,040	0,041	0,040	0,041	0,018
Skew	-0,39	-0,46	-0,37	-0,23	-0,35	-0,38	-0,08	-0,43	-0,28	-0,52	0,28
Kurtosis	1,09	1,15	0,91	0,99	0,83	1,07	0,52	0,49	0,87	0,59	0,48
Sharpe-ratio	0,29	0,23	0,29	0,27	0,31	0,28	0,30	0,37	0,38	0,40	-0,27

Notes:
N: of monthly observations pre-financial crisis: 53
N: of monthly observations post-financial crisis: 102

Table 5.2.1 presents the summary statistics for the two subperiods. The negative effect of ESG ratings on the financial performance is still discernible in the negative monthly average returns of the difference portfolio for both periods. The monotonic trend negative trend for ESG and financial performance is however only visible in the period after the financial crisis. The standard deviation is again congruent across the decile portfolios for both periods.

The Sharpe ratios of the Low-1 portfolio is higher than the High-10 prior to the financial crisis yet neither of the portfolios display an impressive Sharpe ratio compared with the other portfolios. This indicates that the effect prior to the financial crisis is spurious. After the financial crisis the Sharpe ratios display a monotonic trend with ESG having a negative effect on the financial performance. The skewness and kurtosis are significantly different when excluding the financial crisis. There is no discernible trend in the skewness across the portfolios for either of the periods. Nor is there any trend across portfolios for the kurtosis prior to the financial crisis. For the period after financial crisis ESG seem to have a slight negative effect on kurtosis and the trend is monotonic.

Table 5.2.2

Panel A: Monthly average portfolio raw returns and alphas 2003-mid 2007 (Pre-financial crisis)

	High-10	9	8	7	6	5	4	3	2	Low-1	Diff
Average Return	1,22%	1,50%	1,28%	1,50%	1,90%	1,24%	1,53%	1,39%	1,43%	1,68%	-0,46%
	(3.31)***	(4.14)***	(3.10)***	(3.65)***	(4.88)***	(2.94)***	(3.93)***	(3.37)***	(3.30)***	(3.91)***	(1.83)*
CAPM Alpha	-0,01%	0,30%	-0,02%	0,14%	0,65%	-0,07%	0,25%	0,08%	0,03%	0,38%	-0,39%
	(0.07)	(2.02)**	(0.07)	(0.75)	(3.81)***	(0.31)	(1.49)	(0.35)	(0.18)	(1.48)	(1.79)*
FF3 Alpha	0,05%	0,38%	0,12%	0,20%	0,68%	0,15%	0,36%	0,16%	0,02%	0,47%	-0,42%
	(0.24)	(2.27)**	(0.45)	(0.99)	(3.74)***	(0.67)	(1.86)*	(0.53)	(0.09)	(1.55)	(1.65)
Carhart 4 Alpha	0,04%	0,38%	0,11%	0,21%	0,68%	0,14%	0,36%	0,17%	0,02%	0,47%	-0,43%
	(0.23)	(2.21)**	(0.44)	(1.04)	(3.67)***	(0.65)	(1.81)*	(0.55)	(0.10)	(1.57)	(1.69)*

Panel B: Loadings on the on the common risk factors using the Carhart four-factor model

Factors	High-10	9	8	7	6	5	4	3	2	Low-1	Diff
RM	0,941	0,896	0,912	0,961	0,876	0,933	0,857	0,977	1,005	0,911	0,030
	(7.99)***	(9.02)***	(8.37)***	(8.38)***	(7.54)***	(7.33)***	(9.06)***	(9.97)***	(8.86)***	(5.98)***	(0.17)
SMB	-0,103	-0,078	0,049	0,137	0,099	0,008	0,178	0,023	0,111	0,101	-0,204
	(1.17)	(0.88)	(0.33)	(1.06)	(0.74)	(0.07)	(1.58)	(0.21)	(0.73)	(0.66)	(1.34)
HML	-0,108	-0,140	-0,194	-0,093	-0,039	-0,350	-0,151	-0,144	0,033	-0,133	0,026
	(0.82)	(1.12)	(1.02)	(0.77)	(0.28)	(2.63)**	(1.02)	(0.81)	(0.19)	(0.68)	(0.15)
MOM	-0,015	-0,014	-0,060	0,047	0,014	-0,095	0,045	0,078	0,008	0,001	-0,016
	(0.22)	(0.22)	(0.75)	(0.79)	(0.17)	(1.05)	(0.55)	(0.91)	(0.11)	(0.01)	(0.16)

Notes:
T-stat is reported in the brackets
*Asterix denotes * p<0,10, ** p<0,05, *** p< 0,01*
N. of monthly observations 53

Panel A of table 5.2.2 reports the average monthly raw returns of the decile portfolios sorted on ESG ratings and the difference portfolio for the subperiod prior to the financial crisis. The returns of the difference portfolio are negative and significant consistent with the test for the entire period. The portfolio also displays a negative and weakly significant Carhart and CAPM alpha. The best performing portfolio was a mid-range ESG rating portfolio. The results indicate that while the Low-1 portfolio significantly outperforms the High-10 portfolio the alphas are less significant for the subperiod prior to the financial crisis.

Panel B of table 5.2.2 reports the loadings of the portfolios on the common risk factors. The decile portfolios all display a positive and significant loading on the market risk factor, RM. The coefficients are generally smaller than for the entire period. The difference portfolio displays as expected a market beta coefficient close to zero and insignificant. The loadings of the SMB factor display the same monotonically negative trend of ESG ratings on the SMB factor. The loadings are not significant for any of the decile portfolios or the difference portfolio.

This result is consistent with the finding for the entire period that higher ESG ratings are negatively associated with the size premium and tilt towards large cap. The loadings of the HML factor on the decile portfolios are mostly negative with an exception for portfolio two and it is insignificant for all decile portfolios except 5. The difference portfolio displays a positive loading on the factor but unlike for the entire period it is insignificant. The loadings on the MOM factor show unlike for the entire period some portfolios with a positive loading but just like for the entire period the loadings are insignificant.

Table 5.2.3

Panel A: Monthly average portfolio raw returns and alphas mid 2009-2018 (post financial crisis)

	High-10	9	8	7	6	5	4	3	2	Low-1	Diff
Average Return	1,18%	0,93%	1,15%	1,16%	1,33%	1,15%	1,23%	1,53%	1,50%	1,66%	-0,48%
	(2.96)***	(2.32)**	(3.01)***	(2.81)***	(3.14)***	(2.83)***	(3.12)***	(3.75)***	(3.86)***	(4.17)***	(2.78)***
CAPM Alpha	-0,13%	-0,34%	-0,12%	-0,17%	-0,05%	-0,18%	-0,02%	0,21%	0,30%	0,42%	-0,54%
	(0.98)	(2.08)**	(0.72)	(1.02)	(0.27)	(1.17)	(0.14)	(1.18)	(1.66)*	(1.77)*	(2.44)**
FF3 Alpha	-0,12%	-0,31%	-0,11%	-0,12%	0,02%	-0,12%	0,04%	0,28%	0,36%	0,46%	-0,58%
	(0.92)	(1.91)*	(0.69)	(0.76)	(0.10)	(0.81)	(0.28)	(1.59)	(2.11)**	(2.15)**	(3.01)***
Carhart 4 Alpha	-0,05%	-0,23%	-0,04%	-0,06%	0,10%	-0,07%	0,12%	0,34%	0,42%	0,55%	-0,60%
	(0.41)	(1.56)	(0.26)	(0.39)	(0.59)	(0.48)	(0.80)	(2.04)**	(2.66)***	(2.61)**	(3.04)***

Panel B: Loadings on the on the common risk factors using the Carhart four-factor model

Factors	High-10	9	8	7	6	5	4	3	2	Low-1	Diff
RM	0,943	0,897	0,913	0,932	0,939	0,921	0,846	0,899	0,821	0,857	0,086
	(17.53)***	(13.5)***	(16.51)***	(15.02)***	(16.09)***	(17.30)***	(14.18)***	(14.40)***	(12.27)***	(11.17)***	(1.91)*
SMB	-0,020	0,032	-0,009	0,087	0,160	0,156	0,211	0,233	0,287	0,258	-0,278
	(0.27)	(0.38)	(0.14)	(1.07)	(1.63)	(1.75)*	(2.14)**	(2.77)***	(3.19)***	(2.35)**	(4.11)***
HML	0,001	0,028	-0,058	0,021	0,015	-0,010	-0,052	-0,040	-0,171	-0,239	0,240
	(0.01)	(0.31)	(0.81)	(0.22)	(0.16)	(0.11)	(0.55)	(0.56)	(1.55)	(2.56)**	(2.54)**
MOM	-0,215	-0,252	-0,240	-0,205	-0,257	-0,189	-0,245	-0,224	-0,195	-0,278	0,064
	(2.92)***	(2.98)***	(3.68)***	(2.25)**	(3.32)***	(2.55)**	(3.55)***	(3.05)***	(2.11)**	(3.17)***	(1.10)

Notes:
T-stat is reported in the brackets
*Asterix denotes * p<0,10, ** p<0,05, *** p< 0,01*
N. of monthly observations 102

Panel A of table 5.2.3 reports the monthly average raw returns of the decile portfolios sorted on ESG ratings and the difference portfolio for the subperiod after the financial crisis. Just like for the entire sample ESG appears to have a negative effect on returns. In the table the difference portfolio has a negative average monthly return and it is highly significant. The portfolio alphas trend from negative and insignificant for the high ESG portfolios to positive and significant for the low ESG portfolios. This results in the difference portfolio displaying negative and significant alphas for the subperiod.

Panel B of table 5.2.3 reports the loadings of the decile portfolios and difference portfolio on the common risk factors. Unlike for the entire period and the subperiod prior to the financial crisis the difference portfolio displays a weakly significant market risk coefficient. The loadings of the SMB factor display the same monotonically negative relationship between ESG ratings and the SMB factor. The difference portfolio displays a negative and significant loading on the size premium. This result confirms the findings above and it tells us that the size premium is consistent through over the entire period.

The loadings of the HML factor on the decile portfolios display a positive trend with the ESG ratings. The loadings are insignificant for all portfolios except the Low-1 portfolio which has a negative and significant loading on the HML factor. Unlike the period prior to the financial crisis the High-10 portfolio seem to tilt less towards growth stocks while the tilt in the Low-1 portfolio increases. The difference portfolio displays a positive and significant loading on the

HML factor for the period after the financial crisis. The negative and significant loading of the Low-1 in conjunction with its short position in the HML portfolio results in the difference portfolio having a significant and positive loading on the value premium. The loading of the MOM factor is consistently negative and significant for all decile portfolios. The difference portfolio has a positive but insignificant loading on the factor. This indicates that the stocks in the decile portfolio all exhibit poor past returns.

5.3 ESG combined ratings effects on returns 2003-2018

To test if the negative effect is consistent when considering the ESG controversies rating, the test is repeated for portfolios sorted on the ESG combined rating. The idea is that the ESG combined rating penalises companies subjected to negative, significant and widely covered ESG related event.

Table 5.3.1

Panel A: Monthly average portfolio Returns and Alphas 2003-2018 (ESG combined score)

	High-10	9	8	7	6	5	4	3	2	Low-1	Diff
Average Return	0,85%	0,83%	0,87%	0,95%	0,89%	0,93%	0,77%	0,88%	0,83%	1,06%	-0,21%
	(2.27)**	(2.01)**	(2.33)**	(2.60)**	(2.39)**	(2.61)***	(2.08)**	(2.22)**	(2.16)**	(2.87)***	(1.51)
CAPM Alpha	-0,19%	-0,28%	-0,15%	-0,07%	-0,14%	-0,05%	-0,25%	-0,22%	-0,19%	0,05%	-0,24%
	(1.39)	(1.32)	(0.83)	(0.42)	(0.91)	(0.35)	(1.47)	(1.30)	(1.06)	(0.23)	(1.45)
FF3 Alpha	-0,19%	-0,27%	-0,15%	-0,06%	-0,14%	-0,06%	-0,24%	-0,22%	-0,20%	0,04%	-0,24%
	(1.38)	(1.28)	(0.79)	(0.39)	(0.90)	(0.44)	(1.46)	(1.27)	(1.04)	(0.17)	(1.44)
Carhart 4 Alpha	-0,18%	-0,25%	-0,12%	-0,04%	-0,11%	-0,04%	-0,22%	-0,18%	-0,17%	0,06%	-0,23%
	(1.21)	(1.14)	(0.66)	(0.21)	(0.70)	(0.33)	(1.33)	(1.02)	(0.89)	(0.27)	(1.45)

Panel B: Loadings on the Fama-French and Carhart factors

Factors	High-10	9	8	7	6	5	4	3	2	Low-1	Diff
RM	1,076	1,119	1,014	1,027	1,037	1,042	1,043	1,114	1,044	1,048	0,028
	(12.08)***	(9.27)***	(10.67)***	(12.00)***	(10.70)***	(13.74)***	(10.40)***	(10.24)***	(9.17)***	(8.15)***	(0.44)
SMB	0,042	0,081	0,105	0,028	0,069	-0,097	-0,035	-0,001	0,065	0,071	-0,028
	(0.56)	(1.04)	(1.44)	(0.32)	(0.94)	(1.17)	(0.40)	(0.02)	(0.73)	(0.81)	(0.38)
HML	-0,071	0,043	-0,033	0,004	-0,089	-0,122	0,043	-0,020	-0,096	-0,198	0,126
	(0.94)	(0.49)	(0.29)	(0.039)	(0.84)	(1.10)	(0.54)	(0.20)	(0.74)	(2.02)**	(1.97)*
MOM	-0,069	-0,080	-0,094	-0,105	-0,118	-0,054	-0,056	-0,145	-0,094	-0,088	0,018
	(0.95)	(1.41)	(1.46)	(1.91)*	(1.96)*	(0.91)	(0.67)	(2.43)**	(1.43)	(1.87)*	(0.42)

Notes:
T-stat is reported in the brackets
*Asterix denotes * p<0,10, ** p<0,05, *** p< 0,01*
N. of monthly observations 180

Panel A of table 5.3.1 reports the monthly average returns in addition to the alphas for the decile portfolios sorted on the ESG combined rating and the difference portfolio for the period 2003-2018. The returns are positive and significant for the decile portfolios. The difference portfolio still displays a negative return but unlike for the ESG ratings the difference is not significant. The negative monotonic trend where higher ESG resulted in poor financial performance seen

with the ESG rating have disappeared with the ESG combined rating. This result explains the insignificant alpha in the difference portfolio.

Panel B of 5.3.1 reports the loadings of the portfolios on the common risk factors. The difference portfolio show consistency in that the coefficients for the FF3 and CH4 factors display the same sign. The difference portfolio still loads negatively on the SMB factor but unlike for the ESG ratings the loading is insignificant. The positive loading on the HML factor is consistent but it is only weakly significant when using the ESG combined ratings. For the decile portfolios the monotonic trend in the SMB factor that indicated a positive size tilt for higher ESG is no longer discernible. This indicates that the combined ESG ratings corrects for the size factor by larger companies being punished by the controversies rating. This result also indicates that the negative and significant alphas of the difference portfolio in the ESG ratings could be attributed to its loading on the size factor. This result confirms the result in the ESG ratings that ESG is negatively related to the size premium and is correlated with larger market capitalization. This could possibly be explained by larger companies having more resources to allocate towards producing seemingly better ESG disclosures.

The decile portfolios loading on the HML factor displays a more of a monotonic trend than in the ESG ratings. This is evidenced by the difference portfolios significant loading on the factor indicating that the difference portfolio is tilted towards value stocks as oppose to growth stocks. The loadings on the MOM factor show no signs of a trend across portfolios. The Low-1 portfolio loads negatively and significantly on the momentum factor. This indicates that the stocks in the low ESG combined ratings portfolio is possibly affected by some negative ESG related event that is incorporated in the stock price.

5.4 ESG change effect on returns 2003-2018

To test if changes in the ESG have some predictive power on financial performance the stocks are ranked and sorted on their annual absolute change in the ESG ratings. The absolute changes are used as opposed to the relative changes because relative changes would induce bias and result in different loadings for the same absolute change depending on which level of non-financial performance the company had prior to the change. Because of the small dispersion in change across the stocks, quintile portfolios are used. The three middle portfolios show very little difference in change each year, this is fine since the study is interested in testing the effect of the large positive/negative change on financial performance. Table A.8 reports some summary descriptive statistics about the level of change in ESG ratings for the High-5 and Low-

1 portfolio. The portfolios are held for one year and rebalanced at the start of every year with the release of new ESG ratings.

Table 5.4.1

Monthly average portfolio returns 2003-2018 (Portfolios sorted on ESG change)

	High-5	4	3	2	Low-1	Diff
Monthly average return entire period	0,81%	0,85%	0,92%	0,77%	0,92%	-0,11%
	(2.14)**	(2.28)**	(2.52)**	(2.01)**	(2.50)**	(1.21)
Monthly average return pre-financial crisis	1,33%	1,35%	1,60%	1,45%	1,46%	-0,12%
	(3.49)***	(3.59)***	(4.16)***	(3.56)***	(3.92)***	(0.78)
Monthly average return post-financial crisis	1,25%	1,26%	1,27%	1,19%	1,37%	-0,12%
	(3.00)***	(3.17)***	(3.26)***	(3.06)***	(3.56)***	(1.19)

Notes:
T-stat is reported in the brackets
*Asterix denotes * p<0,10, ** p<0,05, *** p< 0,01*
N. of monthly observations: entire period 180, pre-financial crisis 53, post-financial crisis 102

Table 5.4.1 reports the monthly average raw returns of the stocks sorted on ESG change for the entire period, pre- and post-financial crisis. The returns are significant for all the quintile portfolios for all time periods. The negative change portfolio, Low-1 displays a higher average return than the large positive change portfolio, High-5 for all period and subperiods. This results in the difference portfolio having a negative monthly average return, yet the difference is not large enough to be significant.

Table 5.4.2

Panel A: Monthly average portfolio raw returns and alphas 2003-2018 for portfolios sorted on ESG change				Panel B: loadings on the common risk factors using the Carhart four-factor model			
	High-5	Low-1	Diff	Factors	High-5	Low-1	Diff
Monthly average retu	0,81%	0,92%	-0,11%	RM	1,084	1,040	0,044
	(2.14)**	(2.50)*	(1.21)		(11.11)***	(9.56)***	(1.60)
CAPM Alpha	-0,24%	-0,010%	-0,013%	SMB	0,036	0,010	0,026
	(1.54)	(0.64)	(1.60)		(0.50)	(0.15)	(0.64)
FF3 Alpha	-0,24%	-0,010%	-0,013%	HML	-0,071	-0,057	-0,014
	(1.54)	(0.62)	(1.61)		(0.68)	(0.54)	(0.24)
Carhart 4 Alpha	-0,022%	-0,007%	-0,015%	MOM	-0,062	-0,126	0,060
	(1.41)	(0.41)	(1.76)*		(0.81)	(2.33)**	(1.96)*

Notes:
T-stat is reported in the brackets
*Asterix denotes * p<0,10, ** p<0,05, *** p< 0,01*
N. of monthly observations: 180

Table 5.4.2 reports the alphas and the loadings of the High-5, Low-1 and difference portfolios on the common risk factors. All portfolios display negative and insignificant alphas except for

the difference portfolio which displays a negative but weakly significant CH4 alpha. This result is confusing since the alphas are insignificant for the CAPM and FF3. This could indicate that the inclusion of the momentum factor does not add power to the test but rather decreases it and the alpha becomes questionable. The Low-1 portfolio is significantly affected by negative momentum. This result indicates that stocks that are associated with a large negative change in ESG also display a significant negative loading on the momentum factor. The loading of the High-5 portfolio on the momentum is also negative but insignificant.

These results are in line with Scholtens & Zhou (2008) who found that companies that perform socially irresponsible actions tend to do poorly but that the opposite effect was not found for companies who is socially responsible. This indicates that while large a decrease in ESG ratings results in negative momentum while a large increase in ESG ratings do not cause a positive momentum. This in conjunction with the insignificant differences indicates that the changes in ESG, positive or negative have little to no predictive power on the financial performance.

5.5 ESG ratings and performance for economy sectors

Table A.9 reports the summary statistics of the ESG ratings across the TRBC economy group for the stocks on the S&P 500. The mean, max and min value display that there are small differences in the ESG ratings across the sectors. To test if there is a sector-based difference in the effect of ESG ratings on the financial performance and to ensure that previous results are not biased because of tilts towards specific sectors, a double sorted portfolio approach is applied. The companies are sorted on economy sector, and then sorted again on ESG ratings into tercile and 2-quantile portfolios. Tercile and 2-quantile portfolios are used to ensure that the portfolios are sufficiently large. The economy group *telecommunications* was excluded since it only contained three companies. To ensure at least 10 stocks in the high and low portfolios the economy groups *Energy, Basic Materials, Non-cyclical consumer goods and Utilities* use 2-quantile portfolios which split companies at the median ESG rating.

Table 5.5.1

Sector	Energy			Basic Materials			Industrials		
	High	Low	Diff	High-3	Low-1	Diff	High-3	Low-1	Diff
Average Return All	0,81%	1,11%	-0,30%	0,76%	1,09%	-0,33%	0,97%	1,15%	-0,18%
	(1.36)	(1.80)*	(1.59)	(1.50)	(2.10)**	(1.25)	(2.45)**	(2.91)***	(1.16)
Average Return Pre-Financial Cri	2,56%	2,45%	0,11%	1,53%	2,07%	-0,54%	1,56%	1,60%	-0,04%
	(3.05)***	(2.87)***	(0.35)	(3.13)***	(2.90)***	(1.01)	(3.34)***	(3.12)***	(0.13)
Average Return Post Financial Cı	0,50%	0,90%	-0,41%	0,93%	1,24%	-0,32%	1,38%	1,67%	-0,29%
	(0.74)	(1.34)	(1.73)*	(1.63)	(2.11)**	(1.08)	(3.09)***	(4.03)***	(1.83)*
Sector	Cyclical Consumer Goods			Non-Cyclical Consumer Goods			Financials		
	High-3	Low-1	Diff	High	Low	Diff	High-3	Low-1	Diff
Average Return All	0,99%	1,05%	-0,06%	0,60%	0,87%	-0,27%	0,50%	0,67%	-0,17%
	(2.30)**	(2.47)**	(0.31)	(2.68)***	(3.18)***	(1.68)*	(0.89)	(1.64)	(0.81)
Average Return Pre-Financial Cri	1,59%	1,28%	0,31%	0,78%	0,85%	-0,06%	1,38%	1,27%	0,11%
	(3.07)***	(2.16)**	(0.90)	(2.33)**	(2.61)**	(0.25)	(3.17)***	(3.40)***	(0.56)
Average Return Post Financial Cı	1,37%	1,58%	-0,21%	0,93%	1,32%	-0,39%	0,96%	1,31%	-0,35%
	(3.08)***	(3.59)***	(0.91)	(3.67)***	(4.40)***	(2.17)**	(1.67)*	(3.11)***	(1.64)
Sector	Health Care			Technology			Utilites		
	High-3	Low-1	Diff	High-3	Low-1	Diff	High	Low	Diff
Average Return All	0,79%	1,62%	-0,83%	1,00%	1,30%	-0,29%	0,60%	0,60%	0,01%
	(2.37)**	(3.93)***	(3.40)***	(2.25)**	(3.10)***	(1.31)	(2.08)**	(1.77)*	(0.042)
Average Return Pre-Financial Cri	0,75%	2,00%	-1,24%	1,56%	1,76%	-0,20%	1,19%	1,29%	-0,10%
	(1.89)*	(3.54)***	(2.61)**	(2.10)**	(2.49)**	(0.48)	(2.67)**	(2.48)**	(0.39)
Average Return Post Financial Cı	1,36%	1,99%	-0,63%	1,38%	1,72%	-0,34%	0,73%	0,85%	-0,12%
	(3.38)***	(4.37)***	(2.71)***	(2.77)***	(3.80)***	(1.35)	(2.16)**	(2.31)**	(0.90)

Notes:
T-stat is reported in the brackets
*Asterix denotes * p<0,10, ** p<0,05, *** p< 0,01*
N. of monthly observations: entire period 180, pre-financial crisis: 53, post financial crisis: 102

Table 5.5.1 reports the monthly average raw returns for the High, Low and difference portfolios double sorted on economy sector and ESG rating. For the entire period 2003-2018 all sectors except the *utilities* sector display a negative average raw return for the difference portfolio. The difference portfolio is weakly significant for the *non-consumer cyclical goods* sectors and strongly significant for the *Health care* sector. The negative effect of ESG on financial performance in the sample seems to transcend across economy sectors as well. For the period prior to the financial crisis the average raw returns of the difference portfolios are positive for the *energy*, *cyclical consumer goods* and the *financial* sector but the difference is not significant. The other sectors display a negative average raw return for the period with the *health care* sector being the only sector that is significant. For the period post the financial crisis the average raw returns are negative across all sectors. The negative returns of the difference portfolios are significant for the *energy*, *industrials*, *non-cyclical consumer goods* and *health care* sector. The utilities sector is a curious case since it displays negative average raw returns prior to and post financial crisis but positive returns on the entire period. This could possibly be explained by a strong stock performance by high ESG companies during the financial crisis.

Overall the negative effect of ESG on financial performance is consistent across the sectors. It is less obvious when observing the results prior to the financial crisis where three sectors display a positive albeit insignificant average raw return. The results after the financial crisis are expected when figure A.10 is considered. The low scoring ESG portfolios consistently

and significantly outperformed their high scoring counterparts between the years 2010-2015. This result holds across the economy sectors as well as they all display negative returns for the period after the financial crisis.

5.6. ESG long run effect on returns

The movement that spur the development of ESG ratings and investing usually regards humans as the custodians of the planet. The movement stresses the long-term effect, that economic growth should be encouraged with consideration for the long-term effect. Another perspective is that of Friedman (1970) who argue that SRI investing incurs unnecessary costs for shareholders while Barnea & Rubin (2010) show that managers have an incentive to use ESG investing as a tool for personal glorification at the expense of shareholders. As has been shown in previous tests good non-financial performance do not necessarily translate into good financial performance. The results could indicate that ESG incurs costs in the short run but that it translates into a more profitable future. It is therefore of interest to test the long-term effect of a good non-financial performance on the stock performance.

To study the long-term effect of high and low ESG ratings on financial performance companies are ranked and sorted into portfolios based on their ESG rating. The portfolios are held for a five-year period with no rebalancing for the entire holding period. If ESG ratings have a long-term effect on the financial performance the twelve portfolios should show some tendency of over/underperformance.

Table 5.6.1

Portfolio Formation(t) and Holding period(t+5)

t	t+5		High-10	9	8	7	6	5	4	3	2	Low-1	Diff
2003	2008	Monthly average returns	0,99%	1,21%	0,98%	0,98%	1,13%	0,82%	0,77%	0,89%	0,88%	1,02%	-0,04%
			(2.52)**	(3.43)***	(2.21)**	(2.32)**	(2.62)**	(2.04)**	(1.69)*	(2.34)**	(2.15)**	(2.23)**	(0.12)
2004	2009	Monthly average returns	-0,14%	-0,28%	-0,44%	-0,38%	-0,06%	-0,38%	-0,32%	-0,06%	-0,33%	0,06%	-0,20%
			(0.25)	(0.37)	(0.67)	(0.43)	(0.08)	(0.50)	(0.42)	(0.08)	(0.40)	(0.10)	(0.89)
2005	2010	Monthly average returns	0,07%	0,00%	-0,03%	-0,03%	0,09%	0,05%	-0,09%	0,04%	0,37%	-0,06%	0,13%
			(0.10)	(0.00)	(0.04)	(0.03)	(0.09)	(0.06)	(0.11)	(0.04)	(0.38)	(0.07)	(0.48)
2006	2011	Monthly average returns	0,05%	0,04%	0,09%	0,01%	0,18%	0,15%	-0,09%	-0,05%	0,30%	0,32%	-0,27%
			(0.06)	(0.04)	(0.11)	(0.01)	(0.20)	(0.17)	(0.09)	(0.05)	(0.33)	(0.33)	(0.83)
2007	2012	Monthly average returns	0,01%	-0,29%	-0,04%	0,04%	-0,05%	-0,06%	-0,06%	0,41%	-0,39%	0,42%	-0,40%
			(0.01)	(0.30)	(0.05)	(0.05)	(0.06)	(0.06)	(0.06)	(0.46)	(0.36)	(0.41)	(1.07)
2008	2013	Monthly average returns	0,23%	0,15%	0,06%	0,27%	0,16%	0,61%	0,37%	0,74%	0,30%	0,84%	-0,61%
			(0.27)	(0.16)	(0.06)	(0.28)	(0.15)	(0.64)	(0.36)	(0.90)	(0.29)	(0.87)	(1.76)*
2009	2014	Monthly average returns	1,09%	1,21%	1,41%	1,21%	1,53%	1,68%	1,67%	1,72%	1,60%	2,00%	-0,92%
			(1.68)*	(1.49)	(2.29)**	(1.58)	(1.95)*	(2.32)**	(2.11)**	(2.57)**	(2.32)**	(3.56)***	(3.95)***
2010	2015	Monthly average returns	0,98%	0,85%	1,14%	1,29%	1,33%	1,20%	1,36%	1,38%	1,53%	1,60%	-0,63%
			(1.75)*	(1.74)*	(2.26)**	(2.17)**	(2.40)**	(2.12)**	(2.85)***	(2.57)**	(2.82)***	(3.22)***	(3.02)***
2011	2016	Monthly average returns	0,47%	0,42%	0,57%	0,79%	0,74%	0,81%	0,61%	0,87%	1,08%	1,02%	-0,55%
			(0.86)	(0.74)	(1.08)	(1.38)	(1.29)	(1.49)	(1.08)	(1.55)	(2.22)**	(1.77)*	(2.28)**
2012	2017	Monthly average returns	0,76%	0,91%	1,15%	1,02%	0,99%	0,88%	0,89%	1,21%	1,21%	1,27%	-0,51%
			(1.59)	(2.17)**	(2.26)**	(2.26)**	(2.19)**	(1.88)*	(2.05)	(2.57)**	(2.66)**	(2.65)**	(2.52)**
2013	2018	Monthly average returns	1,01%	0,96%	1,11%	1,15%	0,94%	0,86%	1,08%	1,32%	1,07%	1,27%	-0,26%
			(2.35)**	(2.26)**	(2.48)**	(2.72)***	(2.09)**	(2.03)**	(2.66)***	(2.94)***	(2.66)**	(2.70)***	(1.26)
2014	2019	Monthly average returns	0,75%	0,36%	0,45%	0,42%	0,60%	0,50%	0,66%	0,71%	0,77%	0,54%	0,21%
			(1.58)	(0.65)	(0.89)	(0.78)	(1.14)	(0.97)	(1.35)	(1.30)	(1.36)	(0.93)	(0.91)

Notes:
T-stat is reported in the brackets
*Asterix denotes * p<0,10, ** p<0,05, *** p< 0,01*
N. of monthly observations for each portfolio: 75

Table 5.6.1 reports the monthly average raw returns of the ESG sorted portfolios and the difference portfolio with holding periods of five years. Each year decile portfolios and the difference portfolio are constructed, which results in eleven portfolios for twelve subsequent years. The returns of the difference portfolio are negative for ten out twelve holding periods, but it is only significant for five consecutive portfolios.

The period of significant underperformance of the high ESG rated portfolios coincides with portfolios formed during and in the aftermath of the financial crisis. This result indicates that for a period after the financial crisis companies with high ESG significantly underperformed their low ESG rated counterparts, this period can be seen in the figure A.10. The figure displays the monthly average returns for portfolios with holding periods of one year. Since changes in ESG ratings have been shown to have no effect on financial performance in previous tests this should not matter. The figure shows a significant outperformance of the low ESG rated portfolios during the period 2010-2015 which coincides with the significant portfolios in this test.

This result was not significant prior to or during the financial crisis neither is it significant for the last portfolio. This result would suggest that apart from a brief period of consecutive underperformance, there is no significant difference that suggests that ESG ratings have a positive/negative long-term effect on financial performance.

6 Conclusions and further research

This paper aimed to expand the knowledge on the different aspects of ESG ratings and its effect on financial performance. By studying portfolios sorted on ESG ratings with one-year holding periods the period of 2003-2018, the study found that ESG ratings have a negative effect on the financial performance. After risk-adjustment the study found that the difference portfolio still displayed significant alphas, but the study also found that the high and low scoring ESG portfolios differed in portfolio composition with high ESG having a significantly higher tilt towards large cap stocks while low ESG portfolio was tilted towards small cap stocks.

The results were consistent for the period after the sub-prime mortgage crisis but not before where the results were more spurious. For a robustness check the ESG combined ratings were used and the size effect previously seen was mitigated and so was the significant difference in performance. We also tested the predictive capacity of ESG ratings on financial performance by testing if changes in ESG ratings had an effect. The ESG change proved to have no significant effect which suggests that changes in ESG ratings provides no new information and is therefore not priced. This result could indicate that the negative performance seen before could be attributed to style differences in portfolio composition rather than ESG performance. The portfolios sorted on economy sector and ESG rating display the same result of low ESG companies outperforming the high ESG companies and this was especially true after the financial crisis. The long-term portfolios suggest the same thing that after the financial crisis there was a period of very strong performance for stocks with low ESG ratings but that this period only lasted for the period 2010-2015.

It would be interesting to perform the same study again but with a significantly larger sample to get robust industry sector divisions. It would also be interesting to study the effect of the three pillars of E, S & G individually on the financial performance. Previous research has shown that they can have remarkably different results on the performance. Another interesting idea would be to test the cost of ESG investing and if there is significant equal cost for abstaining from ESG investing in the form of reputational cost, non-sustainability risk or litigation risk.

References

Bali, G.T, Engle, F.R, Murray, S. (2016), "Empirical Asset Pricing: The cross section of stock returns". New York: John Wiley & Sons, Inc, Edition 1, pp. 33-88

Barnea, A. & Rubin, A. (2010) "Corporate Social Responsibility as a Conflict between Shareholders" *Journal of Business Ethics 97*, pp. 71-86

Belghitar, Y. Clark, E. Desmukh, N. (2014), "Does it pay to be ethical? Evidence from the FTS4EGood" *Journal of Banking and Finance 47*, pp. 54-62

Buallay, A. (2019), "Is sustainability reporting (ESG) associated with performance? Evidence from the European banking sector" *Management of Environmental Quality: An International Journal 30*, pp. 98-115

Carhart, M.M. (1997), "On Persistence in Mutual Fund Performance" *The Journal of Finance 52*, pp. 57-82

Coram, J.P., Mock, J.T. & Monroe, S.G. (2011), "Financial analysts' evaluation of enhanced disclosure of non-financial performance indicators" *The British Accounting Review 43*, pp. 87-101

Das, N., Chatterje, S., Ruf, B. & Sunder, A. (2018) "ESG Ratings and the Performance of Socially Responsible Mutual Funds: A Panel Study" *Journal of Finance issues 17*, pp. 49-57

Derwall, J., Guenster, N., Bauer, R. & Koedijk, K. (2005), "The Eco-Efficiency Premium Puzzle" *Financial Analyst Journal 61*, pp. 51-63

Donovan, W. (2019) "The Origins of Socially Responsible Investing", The Balance, Viewed 2019-04-10, < https://www.thebalance.com/a-short-history-of-socially-responsible-investing-3025578>

Duuren, V.E., Plantinga, A. & Scholtens, B. (2016) "ESG integration and the investment management process: fundamental investing reinvented", *Journal of Business Ethics 138*, pp. 525-533

El Ghoul, S. & Karoui, A. (2017), "Does corporate social responsibility affect mutual fund performance and flows?", *Journal of Banking and Finance 77*, pp. 53-63

Elzahar, H., Hussainey, K., Mazzi, F. & Tsalavoutas, I. (2015), "Economic consequences of key performance indicators' disclosure quality" *International Review of Financial Analysis 39*, pp. 96-112

Evans, L.J. & Archer, H.S. (1968) "Diversification and the Reduction of Dispersion: An Empirical Analysis" *The Journal of Finance 23*, pp. 761-767

Fama, F.E. (1970), "Efficient Capital Markets: A Review of Theory and Empirical Work" *The Journal of Finance 25*, pp. 383-417

Fama, F.E. & French, R.K. (1992), "The Cross-section of Expected Stock Returns" *The Journal of Finance 47* pp. 426-465

Fama, F.E. & French, R.K. (1993), "Common risk factors in the returns of stocks and bonds" *Journal of Financial Economics 33*, pp. 3-56

Friedman, M. "The social responsibility of businesses is to increase its profits", *The New York Times Magazine*, 1970, pp. 122-126, viewed 2019-05-17

Galema, R., Plantinga, A. & Scholtens, B. (2008) "The stocks at stake: Return and risk in socially responsible investment" *Journal of Banking and Finance 32*, pp. 2646-2654

Hamilton, S., Jo, H. & Statman, M. (1993) "Doing well while doing good? The investment performance of mutual funds", *Financial Analysts Journal 49*, pp. 62-66

Hong, H. & Kacperczyk, M. (2009), "The Price of Sin: the Effect of Social Norms on Market" *Journal of Financial Economics 93*, pp. 15-36

Humphrey, E.J., Lee, D.D & Shen, Y. (2012), "Does it cost to be sustainable?" *Journal of Corporate Finance 18*, pp. 626-639

Ioannou, I. & Serafeim, G. (2017), "The consequences of mandatory corporate sustainability reporting", *Harvard Business School Research Working paper 11-100*,

Jegadeesh, N. & Titman, S. (1993), "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency", *The Journal of Finance 58*, pp. 65-91

Jensen, C.M. (1967) "The Performance of Mutual Funds in the Period 1945-1964", *The Journal of Finance 23*, pp. 389-416

Kempf, A. & Osthoff, P. (2007), "The Effect of Socially Responsible Investing on Portfolio Performance" *European Financial Management 13*, pp. 908-922

Landi, G. and Sciarelli, M (2018) "Towards a more ethical market: the impact of ESG rating on corporate financial performance", *Social Responsibility Journal 15*, pp. 11-27

Limkriangkrai, M., Koh, S. & Durand, B.R. (2017), "Environmental, Social and Governance (ESG) Profiles, Stock Returns, and Financial Policy: Australian Evidence", *The International Review of Finance 17*, pp. 461-471

Manescu, C. (2011) "Stock returns in relation to Environmental, Social and Governance performance: Mispricing or compensation for risk?" *Sustainable development 19*, pp. 95-118

Mollet, C.J. & Ziegler, A. (2014), ”Socially responsible investing and stock performance: New empirical evidence for the US and European stock markets”, *Review of Financial Economics 23*, pp. 208-216

Mossin, J (1966), “Equilibrium in a Capital Asset Market”, *Econometrica: Journal of the Econometric Society 34*, pp. 768-783

MSCI (2018) “MSCI ESG Ratings Methodology”, MSCI ESG Research LLC, Viewed 2019-03-29, <https://www.msci.com/documents/10199/123a2b2b-1395-4aa2-a121-ea14de6d708a>

Refinitiv, (2019), “Thomson Reuters ESG scores”, Refinitiv, Viewed 2019-03-22, <https://www.refinitiv.com/content/dam/marketing/en_us/documents/methodology/esg-scores-methodology.pdf>

Scholtens, B. & Zhou, Y. (2008) “Stakeholder Relations and Financial Performance”, *Sustainable development 16*, pp. 213-232

Sharpe, F, W. (1964) “*Capital Asset Prices: A theory of Market Equilibrium under conditions of risk*”, *The Journal of Finance 7*, pp. 77-91

Sharpe, F, W. (1966) “Mutual fund performance”, *The Journal of Business 39*, pp. 119-138

Statman, M. & Glushkov, D. (2009), “The Wages of Social Responsibility” *Financial Analyst Journal 65*, pp. 33-46

Statman, M. (1987) “How Many Stocks Make a Diversified Portfolio?” *Journal of Financial and Quantitative Analysis 22*, pp. 353-363

Statman, M. (2006) “Socially Responsible Indexes” *The Journal of Portfolio Management 32*, pp. 100-109

SustainAbility (2019) “Sustainability rate the raters”, SustainAbility, Viewed 2019-04-02 <http://sustainability.com/wp-content/uploads/2019/02/sustainability-rate-the-raters-2019.pdf>

UN Global Compact, (2005) “Who cares wins”, UN Global compact, Viewed 2019-04-01, <https://www.unglobalcompact.org/docs/issues_doc/Financial_markets/who_cares_who_wins.pdf>

UNPRI, (2006) “What are the principles for responsible investment”, United Nations Principles for responsible investment, Viewed 2019-04-01 <https://www.unpri.org/pri/what-are-the-principles-for-responsible-investment>

US SIF, (2018) “US SIF Trends report 2018”, US SIF foundation, Viewed 2019-03-05, <https://www.ussif.org/files/US%20SIF%20Trends%20Report%202018%20Release.pdf>

Wesley, J. n.d. “The use of money”, Wesley Center Online, viewed 2019-04-11, <http://wesley.nnu.edu/john-wesley/the-sermons-of-john-wesley-1872-edition/sermon-50-the-use-of-money/>

Appendix

Figure A.1: The percentage of firms on the S&P 500 that obtained an ESG rating

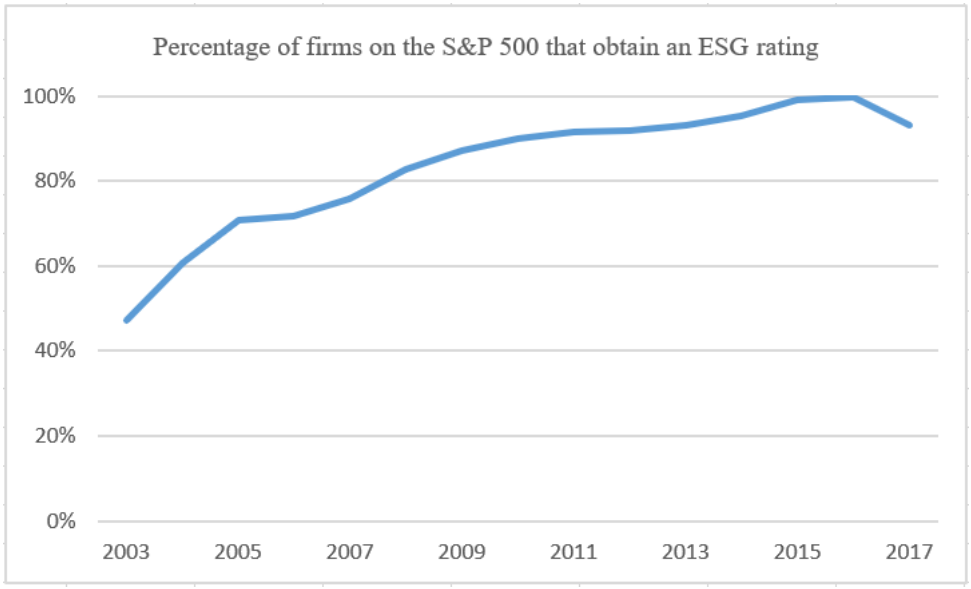

Figure A.2: The 20th, median and 80th percentile of ESG ratings over time.

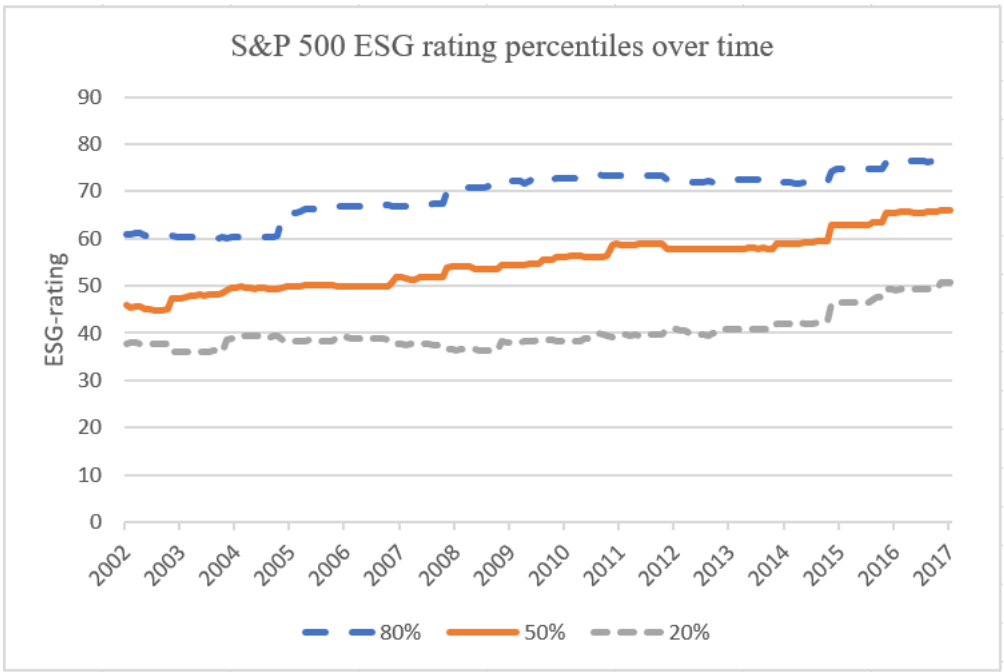

Figure A.3: The number of stocks in each decile portfolio sorted on ESG with the holding periods one year.

	Number of stocks in each portfolio sorted on ESG rating									
	High-10	9	8	7	6	5	4	3	2	Low-1
2003	24	24	22	24	23	23	24	23	24	24
2004	31	30	29	29	30	31	30	30	29	31
2005	36	36	34	34	36	35	35	36	34	35
2006	36	36	36	35	36	35	36	36	35	35
2007	38	38	38	37	38	38	38	37	38	36
2008	42	41	41	41	40	41	43	40	41	41
2009	44	44	43	43	43	43	42	44	43	44
2010	46	44	45	45	44	45	44	45	43	45
2011	46	46	48	42	46	45	46	46	46	45
2012	46	46	46	46	44	46	46	46	45	46
2013	47	47	46	47	47	46	46	46	46	47
2014	48	48	48	48	47	48	47	48	48	46
2015	50	50	50	50	50	49	50	51	49	49
2016	50	50	50	50	50	49	50	50	49	49
2017	47	47	46	47	47	46	47	46	47	47

Figure A.4a: Summary statistics (min, average, max) on the ESG ratings for the companies in each decile portfolio for the period 2003-2017

	High-10			9			8			7			6		
	Min	Average	Max	Min	Average	Max	Min	Average	Max	Min	Average	Max	Min	Average	Max
2003	66,5	74,5	85,8	60,5	63,3	66,3	54,7	57,2	60,4	50,7	52,2	54,3	47,3	48,6	50,6
2004	68,4	76,0	87,5	60,3	63,8	68,3	55,4	57,7	60,1	52,6	53,9	55,4	49,1	51,1	52,4
2005	72,4	80,0	92,4	65,3	68,6	72,4	57,8	61,2	65,1	53,1	55,0	57,6	49,7	51,2	53,1
2006	74,8	79,8	95,7	66,8	70,1	74,3	60,2	63,1	66,8	54,4	57,3	59,9	50,0	51,9	54,1
2007	73,7	80,0	95,2	66,8	69,6	73,2	61,5	63,8	66,8	57,2	59,2	61,5	50,9	54,1	56,9
2008	77,1	82,6	91,9	70,4	73,5	76,7	64,3	67,0	69,9	58,7	61,7	64,2	53,9	56,2	58,7
2009	79,1	84,0	96,5	72,2	75,4	79,0	67,0	69,2	72,1	61,1	64,0	66,9	54,6	57,2	60,3
2010	78,9	84,3	97,5	72,8	75,3	78,9	67,1	69,7	72,6	61,6	64,5	66,8	56,2	59,0	61,4
2011	79,1	84,0	94,1	73,4	75,7	78,9	66,5	69,8	73,3	62,6	64,4	66,3	58,6	60,8	62,5
2012	77,9	82,2	91,2	72,6	75,1	77,8	65,9	69,1	72,5	62,2	63,9	65,8	57,8	59,8	62,2
2013	79,0	83,4	90,5	72,1	75,0	78,8	66,7	69,3	72,1	62,7	64,7	66,6	57,8	60,3	62,4
2014	77,5	82,9	91,7	71,7	74,5	77,5	66,8	69,6	71,6	63,1	64,9	66,6	59,2	61,3	63,1
2015	79,8	84,2	93,0	74,3	77,0	79,7	69,7	72,0	74,2	66,2	67,9	69,7	62,9	64,7	66,1
2016	82,2	85,9	93,3	76,2	78,6	81,9	73,0	74,4	76,0	69,0	70,9	72,9	65,6	67,1	68,9
2017	81,5	85,4	91,2	76,4	78,5	81,2	73,0	74,9	76,3	69,2	70,9	73,0	66,0	67,5	68,8

Figure A.4b Description: same as A.4a

	5			4			3			2			Low-1		
	Min	Average	Max	Min	Average	Max	Min	Average	Max	Min	Average	Max	Min	Average	Max
2003	43,9	45,5	47,3	39,8	41,8	43,8	36,2	37,8	39,7	31,5	34,2	36,1	23,3	27,9	31,5
2004	46,1	47,6	49,0	42,9	44,5	46,0	38,6	40,8	42,8	34,3	36,5	38,4	22,4	30,2	33,9
2005	46,0	48,0	49,6	42,2	43,9	45,9	38,6	40,4	42,0	34,0	36,2	38,4	22,4	30,0	33,4
2006	46,1	47,7	50,0	42,6	44,4	46,1	39,1	40,9	42,4	34,0	36,8	39,0	15,7	28,8	33,8
2007	46,6	48,6	50,8	42,7	44,5	46,5	37,7	39,9	42,6	32,7	35,3	37,7	16,1	27,0	32,6
2008	48,5	51,1	53,7	42,4	45,0	48,4	36,7	39,5	42,3	31,8	34,4	36,6	15,4	26,6	31,6
2009	48,4	51,5	54,4	43,2	45,9	48,1	38,4	40,8	43,2	31,1	34,2	38,0	11,4	25,8	31,0
2010	50,5	53,3	56,2	44,7	47,3	50,5	38,3	41,7	44,6	31,8	35,7	38,3	13,3	26,8	31,8
2011	52,8	55,9	58,4	45,7	49,7	52,6	39,2	42,4	45,4	32,4	35,6	39,1	14,9	24,9	32,1
2012	53,2	55,5	57,8	47,5	50,3	53,1	40,8	43,7	47,4	32,6	36,2	40,6	8,0	25,7	31,8
2013	53,9	55,9	57,6	48,3	51,7	53,9	40,9	44,4	48,1	32,7	37,2	40,8	12,0	26,2	32,3
2014	54,6	56,7	58,9	48,6	51,7	54,5	42,2	45,3	48,0	33,7	38,5	42,0	13,7	26,4	33,6
2015	58,6	60,9	62,9	53,2	56,0	58,5	46,3	49,7	53,0	37,0	41,2	46,2	19,6	30,3	36,8
2016	61,3	63,5	65,5	56,4	59,2	61,3	49,2	52,4	56,2	39,9	45,1	49,0	23,4	32,9	39,8
2017	62,0	64,2	66,0	56,9	59,4	61,9	50,9	53,9	56,7	42,4	47,5	50,7	21,6	35,0	42,4

Figure A.5

The average monthly returns for the entire period 2003-2018 for the companies sorted on ESG rating.

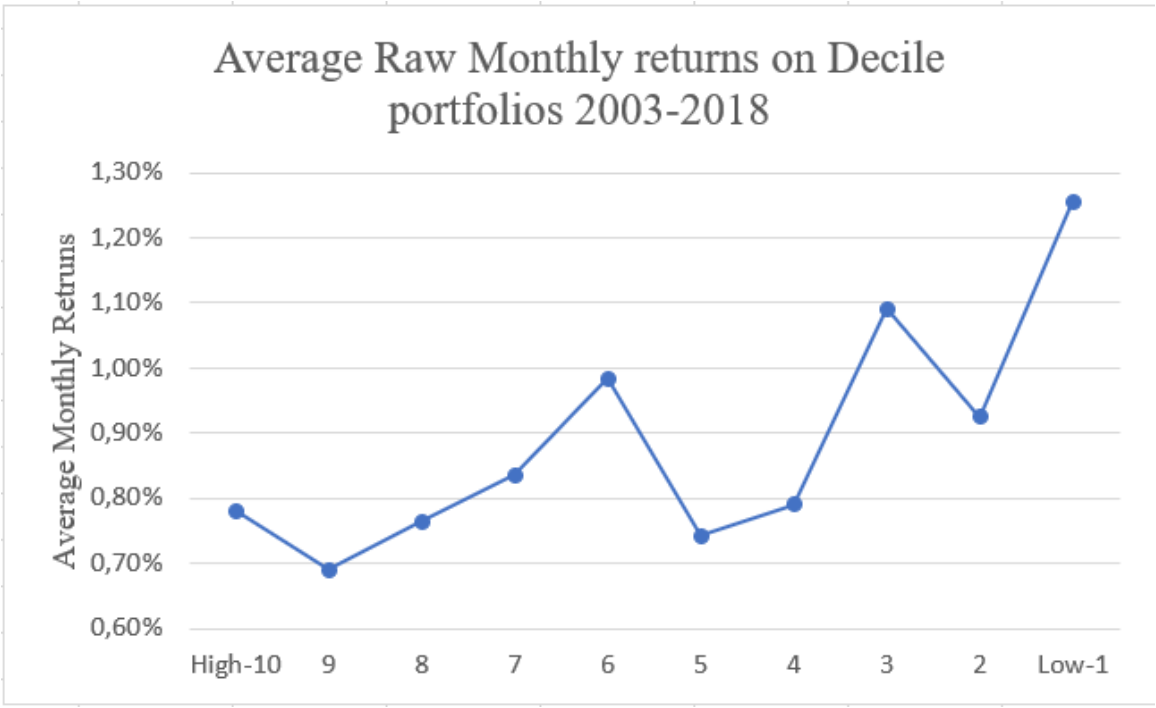

Figure A.6

The monthly returns of the difference portfolio for the stocks sorted on the ESG rating for the period 2003-2018.

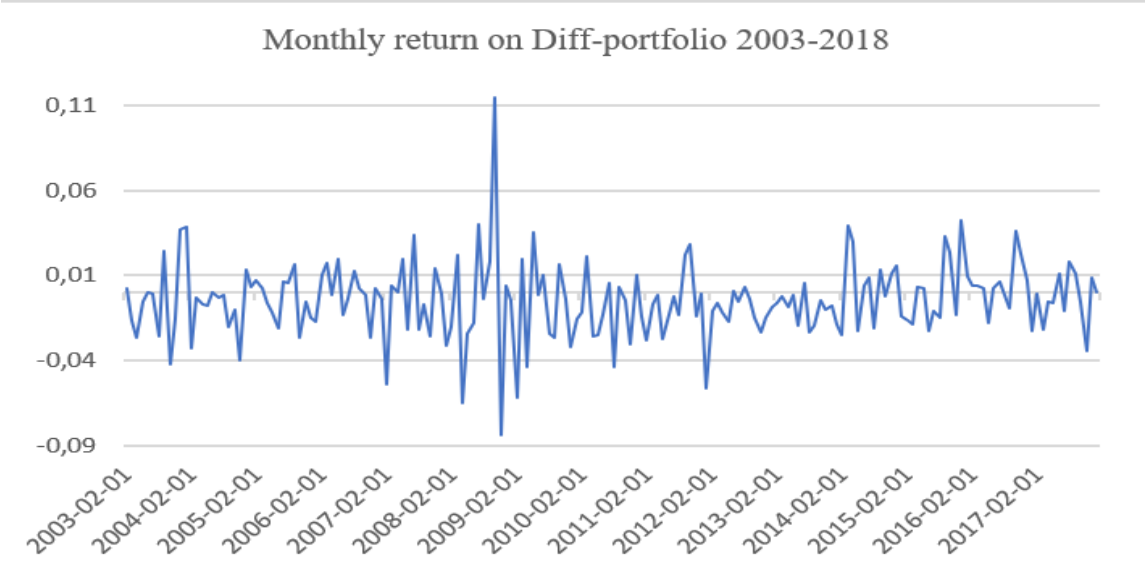

Figure A.7

The development on a 1$ investment in the difference portfolio for the stocks sorted on ESG ratings for the period 2003-2018.

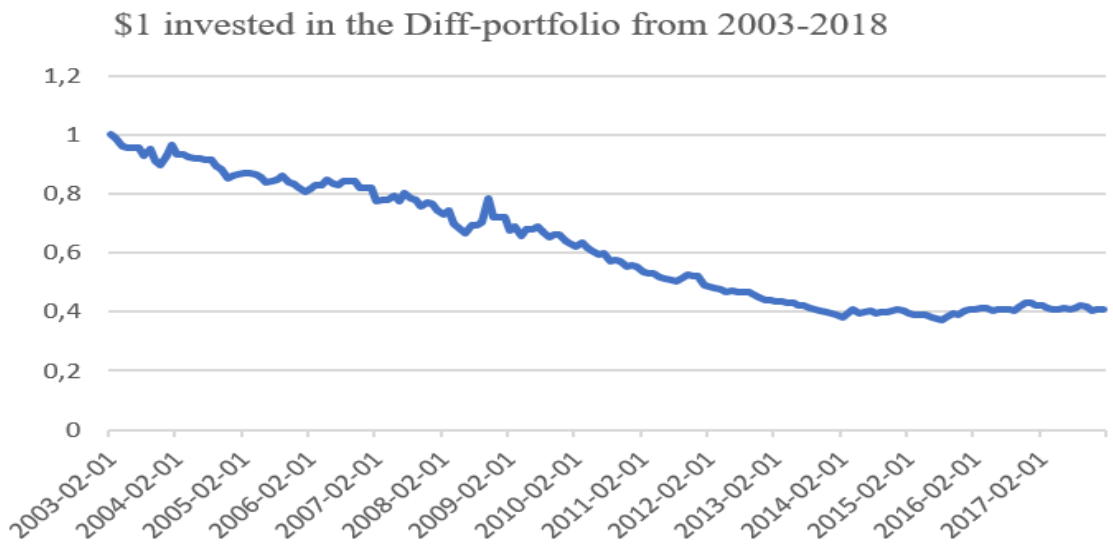

Table A.8:

Summary ESG rating statistics for portfolios sorted on ESG change
The table reports the ESG statistics about the portfolios with largest and smallest change

	High-5			Low-1			High-5	Low-1
	Min	Average	Max	Min	Average	Max	N. of stocks	N. of stocks
2003	3,4	11,1	27,1	-17,2	-9,3	-5,4	48	47
2004	9,8	13,8	27,2	-22,7	-8,0	-3,9	48	48
2005	7,9	15,6	35,5	-28,8	-11,1	-5,8	61	61
2006	6,7	13,6	43,0	-30,8	-10,6	-5,9	71	71
2007	9,2	18,2	41,4	-25,6	-12,6	-8,5	72	72
2008	9,0	13,5	31,6	-15,9	-6,7	-3,2	76	76
2009	7,0	11,5	21,0	-16,7	-6,8	-2,6	83	83
2010	6,4	12,2	26,3	-25,7	-7,5	-4,2	88	87
2011	5,9	10,9	33,0	-12,6	-7,0	-4,2	91	90
2012	3,5	9,3	26,6	-74,6	-7,7	-4,4	92	92
2013	3,8	7,5	63,7	-13,7	-4,8	-2,5	92	92
2014	5,5	9,9	26,7	-32,5	-7,5	-3,7	94	93
2015	8,1	12,6	28,6	-15,3	-3,9	-0,5	96	96
2016	6,9	10,9	42,4	-14,9	-4,3	-1,6	100	99
2017	4,5	9,4	26,3	-12,3	-5,5	-2,9	93	93

Table A.9: Summary statistics (min, average, max) on the ESG ratings for the portfolios sorted on Thomson Reuters Economy sectors for the period 2003-2017.

	Basic Materials			Cyclical Consumer Goods			Non cyclical consumer goods			Energy			Financial		
	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max
2003	23,3	47,9	85,8	23,7	48,3	83,9	26,7	45,6	69,9	24,0	53,1	78,3	25,5	46,2	67,5
2004	22,4	49,9	87,5	28,2	49,2	85,5	22,4	47,0	78,9	30,4	54,6	87,5	24,5	48,1	72,5
2005	22,4	51,2	92,4	27,5	48,7	80,3	26,4	51,4	92,4	32,0	54,9	88,9	28,5	48,0	81,7
2006	15,7	51,8	95,7	15,7	48,2	81,1	22,2	53,7	95,7	28,7	54,5	82,6	27,4	49,8	78,3
2007	16,1	52,0	95,2	16,1	48,6	95,2	22,3	55,0	85,4	27,0	52,1	88,2	23,6	49,7	85,0
2008	15,4	53,8	91,9	15,4	51,0	88,6	25,0	57,9	84,6	28,2	55,5	91,9	21,3	51,8	86,2
2009	11,4	54,9	96,5	11,4	52,2	84,3	22,9	59,6	86,9	22,0	56,2	87,6	22,5	52,8	87,0
2010	16,8	55,8	97,5	23,0	53,6	91,3	22,8	62,0	89,5	20,9	56,8	89,5	24,7	52,3	86,6
2011	15,8	56,4	94,1	17,8	52,4	90,6	18,6	62,6	90,7	19,7	58,2	86,7	19,0	54,8	86,9
2012	8,0	56,2	91,2	19,6	52,6	87,6	21,6	63,9	87,4	17,6	57,1	87,4	20,4	54,4	86,3
2013	12,0	56,9	90,5	22,0	52,3	90,5	20,4	64,9	84,6	17,4	57,4	90,1	20,8	55,4	88,3
2014	13,7	57,2	91,7	20,6	53,6	90,2	20,8	64,1	81,8	20,0	58,8	91,7	19,9	55,2	88,4
2015	19,6	60,4	93,0	24,3	56,9	89,1	20,5	62,3	83,2	29,3	60,5	84,7	23,5	59,9	88,8
2016	23,4	63,0	93,3	26,5	59,3	90,7	27,3	66,4	86,0	32,9	61,6	88,5	23,4	61,9	91,7
2017	21,6	63,7	91,2	21,6	61,4	91,2	29,8	66,0	88,3	35,7	63,9	89,7	22,7	63,2	89,5
	N	23		N	81		N	37		N	29		N	98	

	Healthcare			Industrials			Technology			Telecommunications			Utilities		
	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max
2003	35,6	54,7	85,8	23,3	45,0	82,0	30,3	52,5	82,2	30,3	52,5	82,2	26,6	51,3	78,8
2004	34,3	55,4	85,2	24,0	48,8	84,4	26,3	53,4	80,0	26,3	53,4	80,0	27,1	54,2	80,6
2005	27,7	54,4	83,7	22,4	51,4	83,9	25,5	54,9	85,0	25,5	54,9	85,0	26,9	53,1	73,9
2006	31,3	54,6	84,9	27,3	52,2	84,6	25,0	55,4	90,5	25,0	55,4	90,5	29,5	54,0	79,0
2007	24,8	53,4	81,6	25,2	50,8	85,3	19,3	55,9	88,7	19,3	55,9	88,7	33,2	56,5	73,2
2008	26,7	53,4	87,9	18,6	51,0	88,0	21,5	55,9	87,6	21,5	55,9	87,6	41,7	61,2	80,2
2009	24,9	56,5	90,6	18,0	50,8	94,2	26,0	56,2	96,5	26,0	56,2	96,5	31,4	61,2	84,9
2010	24,3	56,3	88,2	13,3	53,1	90,9	20,5	58,5	97,5	20,5	58,5	97,5	29,8	60,0	84,1
2011	20,7	55,9	84,4	14,9	54,1	90,9	15,8	58,3	94,1	15,8	58,3	94,1	30,1	60,6	80,9
2012	23,8	56,3	85,9	8,0	52,4	87,6	13,7	58,3	91,2	13,7	58,3	91,2	35,4	59,7	80,7
2013	27,5	57,3	86,0	13,4	54,0	87,0	12,0	58,9	90,1	12,0	58,9	90,1	38,0	60,2	84,3
2014	22,9	57,7	86,1	13,7	55,3	89,6	18,5	58,8	91,2	18,5	58,8	91,2	24,1	60,0	79,3
2015	26,5	61,5	89,7	19,6	58,5	91,9	22,8	62,5	93,0	22,8	62,5	93,0	28,9	64,0	81,1
2016	32,2	65,2	87,8	24,3	61,3	87,3	24,7	65,7	93,3	24,7	65,7	93,3	28,1	65,7	83,5
2017	31,6	66,0	86,8	27,6	60,8	89,7	29,8	65,2	90,6	29,8	65,2	90,6	31,4	64,9	83,3
	N	58		N	75		N	69		N	3		N	28	

Figure A.10

The average monthly return for the High-10 and Low-1 portfolio for each year for the period 2003-2018