Erasmus School of Economics

Master Thesis Economics & Business

Financial Economics

Financial Performance and Utility Function of Socially Responsible Investors

A research on the effects of sustainability ratings on the returns of mutual funds, the returns of stocks and the stock price reactions on earnings announcements.

Author: Sami Gamil

Student ID: 482618

Supervisor: Dr. Marshall X. Ma

Second assessor: Dr. Antti Yang

Date final version: 22-11-2022

Erasmus University Rotterdam
Making Minds Matter

Table of contents

Abstract

This study researches the effect of Socially Responsible Investing (SRI) on returns and the utility function of Socially Responsible (SR) investors. I used cross-sectional and panel data from Morningstar and Refinitiv to extract sustainability ratings for mutual funds and firms' ESG scores, respectively. Three hypotheses have been tested using OLS regressions, portfolio sorting and an event study. I find that SR funds have a significantly lower annual return than conventional funds in 2021, where a lower rating on a scale of 1-5 decreases the annual return by around 1%. Also, high ESG stocks are significantly outperformed by low ESG stocks in the period 2002-2020. A portfolio with a long position in high ESG stocks and a short position in low ESG stocks generates significant negative abnormal returns, varying between 0.4% and 0.5% a month depending on the quantile and factor model. Lastly, I do not find a significant difference in stock price reactions to earnings surprises for high and low ESG stocks in the third quarter of 2021, except for a 7-day event window, where low ESG stock prices tend to react more heavily on positive earnings surprises, with a difference of 1.95%.

1. Introduction

More than ever, the world population is concerned with environmental and social problems. The news is filled with items on rising sea levels, global warming, and gender and racial inequality. Individuals and companies are increasingly aware of the problems and are undertaking measures. Decreasing the carbon footprint or equal pay between males and females are two common measures. For companies, this kind of behavior is called Corporate Social Responsibility (CSR). CSR is a broad term that includes behaviors such as being employee- and environment-friendly, ethical, and investor-friendly (Bénabou & Tirole, 2010). CSR merges social goals with financial goals. Companies such as Starbucks are known for their ability to combine financial goals with social goals in terms of keeping their investors happy. Starbucks does this by keeping high social standards, such as policies for human labor; no child labor and respectable wages. Starbucks also only uses sustainable palm oil as one of its environmental measures (Committed to Transparency—People, Planet, Coffee. n.d.). Being socially responsible is also a popular choice for individual investors and funds. This is called Socially Responsible Investing (SRI). More and more investors are incorporating externalities in their investment choices, such as social and ethical considerations (Cortez et al., 2012; Renneboog et al., 2008). Socially responsible (SR) investors base their investments not only on the standard investment criteria, they add criteria based on societal, ethical, and governance issues. Based on these criteria, these investors decide whether to include or exclude certain assets according to Renneboog et al. (2008). Besides private investors, also institutions such as mutual funds have included a firm's social performance in their investment criteria (Cao et al., 2020). More specifically, institutions look at ESG scores. Which stands for Environment, Social, and Governance scores. Firms with high ESG scores tend to do better in these subjects.

However, the main goal of funds and individual investors is to achieve high returns. SR investing and corresponding returns have been researched in the past. Is SRI as profitable as conventional investing? The theory is conflicting regarding the answer to this question (Awaysheh et al., 2020). Therefore, the main interest in this research will be whether SRI is as profitable as conventional investing and whether (SR) investors have a utility function that not only depends on optimizing the risk-reward ratio, which is assumed by Bollen (2007). The main research question is constructed accordingly:

Does Socially Responsible Investing affect the investor's potential return and utility function?

This research question will be answered through three sub-questions, which are as follows.

- *Do Conventional funds differ in performance from Socially Responsible funds?*
- *Does an ESG factor help in explaining the returns of stocks?*

- *Are socially responsible investors less sensitive to earnings surprises than conventional investors?*

The first sub-question answers the question of whether SR funds generate the same returns as conventional funds. If this is the case, it is proof that SRI and returns can go hand in hand. The second sub-question is about the ESG factor. Previous literature has shown that several factors generate significant positive excess returns. Such as size, book-to-market value, and momentum (Fama & Macbeth, 1973; Carhart, 1997). I will test whether an ESG factor is also able to generate positive returns. Positive returns would then indicate that SRI is profitable and taking ESG -scores in mind will increase excess returns. The last sub-question is more focused on the behavior of SR investors. Bollen (2007) assumes that SR investors' utility function not only depends on the risk-return ratio. Instead, it could include social responsibility and conscience. To find out, I will research the sensitivity of conventional and SR investors to earnings surprises. If SR investors are less sensitive to this matter, it means that SR investors care more about SRI than the pure risk-return ratio.

This research would contribute to both society and science in several ways. There have seen several types of research on the topic of ESG investing and whether adding ESG criteria increases the returns of the investments. However, most of these researches have been conducted using data until the year 2013 or even older. Halbritter & Dorfleitner (2015) discussed that the SRI market is continuously developing. Therefore a current sample is critical. This article is written in 2015, and ever since then, ESG investing has become significantly more popular. This would be one of the first researches to use current data. Even more, the conclusion of these researches is conflicting. There is no clear answer to this problem, which can be solved by this research using different methods: Finding out whether ESG funds are outperformed by conventional funds and finding out whether an ESG factor generates significant positive returns. These methods will be further explained in section 3. Therefore, ideally, this research would answer the existing question of whether SRI is possible without losing profit.

Furthermore, this research will answer the question of whether SRI investors differ from conventional investors. In the sense that investors are willing to take a lower profit in exchange for 'doing good' to society. This would mean that the investors are not rationally based on the general utility function that solely depends on the risk-return ratio. Another conclusion could be that the general utility function does not solely depend on the risk-return ratio. This question will be answered by the third sub-question. The method for answering this question will be explained in section 3.

In general, this research will teach us about the benefits of ESG investing, the opportunity costs, and the mindset of SR investors. Especially in a time like today, where the negative externalities of 'sin stocks' cannot be neglected, it is important to learn about ESG investing and the behavior of investors.

The paper is structured as follows; first, the theoretical framework describes previous literature regarding this subject. It also contains theory on the subjects of matter and explains the terms used in this thesis. After that, the research methods and the data selection are described in the section Methods & Data. The research methods are then applied to the data to find an answer to the hypothesis. The outcomes are shown in the Results section. Finally, the Conclusion & Discussion section describes whether and why the hypotheses are rejected or accepted. Imperfections and points of improvement of this study are discussed and applied to provide advice for future related studies.

2. Theoretical Framework

The theoretical framework will be the basis of the theory behind the three sub-questions. First, the previous literature on the subject of this thesis will be summarized and discussed. After that, the relevant theory is discussed.

2.1 Previous Literature

To gain insight into the returns of SR and conventional funds. The papers that researched differences in the returns between those types of funds are discussed. Next, research on an ESG factor is discussed, as well as research on other SR factors. This is to gain insight into whether an ESG factor generates significant excess returns. Lastly, papers on the utility function of (SR) investors are reviewed. The utility function means the preference set of these investors, therefore, literature regarding empirical research on the preferences of investors, and whether SR investors care about more than solely the risk-return ratio. Based on the findings in the previous literature, hypotheses are elucidated for each of the three sub-questions.

2.1.1 Return differences between SR and conventional funds

Socially Responsible funds have been compared to conventional funds in the past. Cortez et al. (2012) did research on global socially responsible funds. More specifically, European and US global socially responsible firms. The authors combine several measures to estimate the funds' outperformance: Jensen's alpha, Christopherson's approach, and a multi-factor model. This leads to a time-varying multi-factor model that estimates Jensen's alpha. The dataset contains 39 funds in European markets and 7 US funds. The period is from 1996 to 2008. The results are in line with other literature according to the authors, socially responsible funds do not significantly outperform conventional funds and benchmarks. Moreover, Austrian and US funds even show negative alpha's.

A more recent paper, written by Durán-Santomil et al. (2019), has researched the fund performance of Sustainable funds. The authors compared mutual funds by their self-proclaimed socially conscious character, as well as their sustainability score from Morningstar. Both of these measures are regressed on the alpha, return, and Sharpe ratio of the funds, along with several control variables. The data used ranges from 2016 to 2018, since Morningstar sustainability ratings are only available from 2016. The funds researched are 1690 all European equity funds that have been scored by Morningstar. The results showed that both the socially conscious and Sustainability score had a positive effect on all of the dependent variables, which are alpha, return, and Sharpe ratio. This research concludes that both of the Sustainability variables can help explain the returns of these funds. This paper also mentions

that there is a lack of clearly defined criteria to rank funds based on SRI. Different measures can yield different results.

Several less recent papers have researched the SRI performance amongst funds. Since these researches are somewhat outdated, I will briefly mention their findings. Most of these researches do not find a significant difference between conventional funds and SR funds. The papers that did not find any significant alpha between conventional funds and SR funds are Hamilton et al. (1993), Mallin et al. (1995), Gregory et al. (1997), and Kreander et al. (2005). All of the above papers estimated alpha's and Jensen's alpha. All of these papers found minimal differences between conventional and SR funds, but none of them were statistically insignificant. Another study by Bauer et al. (2005) found different results for different regions, where ethical US funds underperformed conventional funds. However, ethical UK funds outperformed conventional funds. In conclusion, the older papers do not seem to find any differences between SR and conventional funds. This could be mainly due to the period in which research was conducted, SRI was not very big and popular during that time. Therefore, there were few SR funds available for research, making it difficult to find significant differences.

2.1.2 A socially responsible factor

One of the earlier researches is done by Derwall et al. (2005), who find that SRI generates superior portfolio performance. This is based on the Innovest Strategic Value Advisors' eco-efficiency scores, where firms are scored on the amount of waste generated against their total production output. The authors merge this data with data gathered from the CRSP stock database. The authors construct two portfolios, a portfolio with high-ranked companies and a portfolio with low-ranked companies. Even after controlling the CAPM model (Sharpe, 1964) and the Carhartt (1997) four-factor model, SRI generates higher returns than conventional investing for a sample from 1995 to 2003. Also after controlling for different types of transaction costs and industries, the results were significant.

This research is backed up by several other types of research showing that SRI generates abnormal returns. Kempf & Osthoff (2007) use a long-short strategy to demonstrate the abnormal returns SRI generates. The authors go long in stocks with the top 10% highest socially responsible rating and go short in stocks with the bottom 10% socially responsible rating. This generates up to 8.7% abnormal returns per year. The authors use the Carhartt (1997) four-factor model, to measure the performance of the portfolios. These returns remain significant after controlling for transaction costs. This research has gathered data from the KLD Research & Analytics from 1992 to 2004.

Several other papers find the opposite results. Belghitar et al. (2014) find no evidence of the outperformance of high ESG portfolios compared to conventional portfolios. The authors even state that a portfolio that goes long in conventional stocks and short in ESG stocks shows positive abnormal

returns. The data used in this research is from the FTSE4good series from 2001 to 2010. Another research, done by Hong and Kacperczyk (2009), investigates sin stocks. Sin stocks are firms in industries such as alcohol, tobacco, gambling, and weapons. Sin stocks tend to outperform compared to conventional stocks. This is according to the hypothesis of the authors. They state that investors pay a price to abstain from investing in these stocks. This leads to the conclusion that it is not profitable to be an SR investor. The authors used a dataset from 1965 to 2006 and controlled for several other factors, such as size and book-to-market ratio.

On top of that, Pedersen et al. (2021) created a model with three types of investors. ESG-unaware, ESG aware and ESG-motivated investors. They show that the ideal portfolio is on the ESG-efficient frontier, which is the highest ESG score for a given Sharpe ratio. However, ESG-motivated investors demand a minimum level of ESG scores, which ultimately lowers the maximum achievable Sharpe ratio, meaning that ESG-motivated investors, in this case, the SR investors, have a lower risk-adjusted return on their portfolio.

Lastly, Cornell (2021) states that an ESG factor does not generate positive abnormal returns, due to the stocks being overpriced. The popularity of high ESG companies makes the price rise of the stock. A definite conclusion is still not made. The ESG ratings are often controversial and contradicting and the ESG data is still very short.

In conclusion, literature has shown different outcomes to the dilemma of whether there is a socially responsible factor. A study that has recently been published (Lioui & Tarelli, 2022), has specifically researched the ESG factor. This paper states that the alpha of the ESG factor has time variation and is dependent on the risk aversion of investors and their desire for ESG. Also, rating agencies give different scores to firms, leading to implications for finding the actual performance of an ESG factor. Therefore, to find the performance of the ESG factor the disagreement of rating firms should be taken into account, as well as time variation. The discussed papers show different results, this could be due to the usage of different data sources, and different periods.

2.1.3 Utility function of SR investors

This section describes literature that has researched the utility function of SR investors, more generally, the behavior of SR investors. The theory of the utility function of investors is based on the paper by Bollen (2007), who claims that investors derive utility from being socially responsible. This paper is one of the first papers that deeply examine the behavior of SR investors. The first paper to find evidence of different behavior was from Geczy et al. (2003), who found that SR investors were more loyal to their fund, by stating that SR investors withdrew capital at a slower rate. Bollen (2007) also found somewhat similar evidence. Fund flows of SR funds have significantly lower monthly volatility than

conventional funds. These findings are leading to conclusions on the utility functions of SR investors, which do not solely depend on the risk-return ratio, but also on an SR attribute. Bollen (2007) used U.S. Mutual fund data from CRSP between 1961-2002. The author split the sample into a conventional funds and SR funds group. The results showed that the volatility of cash flows in the SR funds was significantly lower than for conventional funds. Even more, SR investors respond more heavily to positive returns and have a lesser response to negative returns. Meaning that SR investors tend to derive utility from the SR aspect.

An interesting paper did unique research on investor behavior (Riedl & Smeets, 2017). The Authors combined empirical research with an experiment. Data is retrieved from one of the largest Dutch mutual fund providers. The data ranges from 2006 to 2012 and contains administrative data of each individual investor. An investor is socially responsible when it holds at least one SRI fund in its portfolio. A survey and experiments were also conducted. The survey contained questions on investment goals, social goals, and social preferences. The experiments were designed to highlight the investors' risk preferences and social preferences. The authors find several interesting results. First of all, intrinsic social preferences impact the decision to invest socially responsibly, investors who acted altruistically in the experiment, were more likely to hold SRI funds. Even more, SR investors are more likely to donate to charity. Investors are willing to pay a higher management fee for SRI funds and expect SRI funds to underperform relative to conventional funds. Overall, the results indicate that investors are less focused on the highest returns, but on social preferences. In light of Bollen (2007), this means that the utility of SR investors is dependent on their social preferences.

Lastly, the paper of Renneboog et al. (2011) studies the money flows in SRI funds. Their findings are in accordance with the findings of Bollen (2007), SR investors are less sensitive to past negative returns of mutual funds. Data from CRSP US mutual fund Database is used for the US funds and S&P Fund Service is used for the funds from Europe, Asia, and Africa. The paper categorized the types of SR funds by the types and categories of SRI screens used. There are negative and positive screens in the categories sin, ethical, social, and environmental. The results show that SR investors care less about negative past returns based on the fund outflows than conventional investors. This is especially the case for negative screens based on sin and ethical issues, meaning that these funds omit certain stocks from their portfolio. Sensitivity to past positive returns is also dependent on the types of screens. Social screens show a weak relationship between positive returns and inflows, however, Environmental screens show a much stronger relationship. The conclusion of this paper is in line with Bollen (2007) and shows that SR investors derive nonfinancial utility from SRI.

2.2 Hypotheses

Based on the previous literature, three hypotheses will be formed for each sub-question. First of all, when comparing the performance of SR funds to conventional funds, there are a lot of different approaches. The definition of an SR fund is not clearly defined (Durán-Santomil et al., 2019). Therefore, different definitions of what an SR fund is yields different results. The paper itself finds that European SR funds yield higher returns than conventional funds, even after adjusting for risk. However, other papers find contradicting evidence or do not find any significant differences between the types of funds. In conclusion, there is no definite evidence of any differences in performance between SR funds and conventional funds. Therefore, my first hypothesis is as follows:

Hypothesis 1: The performance of Socially Responsible funds does not significantly differ from conventional funds.

The second sub-question is about the ESG factor. Theory regarding this topic is not all in line with each other. Several papers have shown that SR stocks significantly outperform conventional stocks, even in multi-factor models (Derwall et al., 2005; Kempf & Osthoff, 2007). On the other hand, there are papers that have stated the opposite (Belghitar et al.,2014; Hong and Kacperczyk, 2009). Finally, based on the model of Pedersen et al. (2021), the efficient portfolio is on the ESG-frontier, meaning that ESG-aware investors are able to create a more efficient portfolio than ESG-unaware investors. Based on this theory, adding ESG scores into a multi-factor model could explain the returns of stocks and portfolios. The second hypothesis states:

Hypothesis 2: An ESG factor helps explain abnormal returns in a (multi)-factor model.

Thirdly, the utility function of SR investors is examined. The utility function is introduced by Bollen (2007), who examines where SR investors derive utility from. He finds that SR investors derive utility from nonfinancial matters by looking at funds inflows and outflows. SR investors are less sensitive to negative past returns. This is backed up by Renneboog et al. (2011) and Riedl & Smeets (2017). However, SR investors seem to be more sensitive to positive past returns. The theory seems to be agreeing with each other, thus I will hypothesize the same results while keeping an open mind. Therefore the third sub-question will have two hypotheses:

Hypothesis 3a: Socially Responsible investors are less sensitive to a negative earnings surprise than conventional investors.

Hypothesis 3b: Socially Responsible investors are more sensitive to a positive earnings surprise than conventional investors.

Even though I will be testing the SR investors' sensitivity differently than in previous literature, I believe the behavior will be revealed in a similar manner.

2.3 Research Context

This section will provide important information on the topic of Socially Responsible Investing. First, the history of SRI will be summarized. How has SRI developed over the years, what do we see as a socially responsible investment, and what are general opinions on SRI? After that, the different ways to measure SRI will be discussed. There is more than one option and none of them is not controversial.

2.3.1 History of SRI

The main thought of SRI is that investments should not be only about the profit, but also about what happens with the investment and to whom does the money go to. ''Any individual or group which truly cares about ethical, moral, religious or political principles should in theory at least want to invest their money following their principles'' (Miller, 1992, p. 248). The first SR investors were church investment bodies, Only since the 1970s, SRI has become an investment goal for certain publicly offered funds. The funds were focused on excluding certain firms that were considered harmful to society or unethical (Sparkes & Cowton, 2004). The typical industries these companies were in are alcohol, tobacco, and weaponry for example. The classical idea that you should invest in following your principles is in line with this first definition of SRI. Excluding stocks that you do not approve of. It became clear that it was not possible to totally exclude all companies with controversial attributes. Funds had a certain threshold, meaning a maximum percentage of the companies' turnover in controversial areas (Rockness & Williams, 1988). This resulted in criticism because people would feel that excluding 'bad' companies would fulfill their ethical obligations. Also, good companies would receive too little attention because of the negative screening. Therefore, positive screens were now implemented, to invest more in companies with good working conditions, no discrimination, and are environment friendly (Bourke, 1997).

With the rise of SRI, Corporate Social Responsibility (CSR) also came up. At the end of the 90th century, the amount and size of SRI funds increased drastically (McCann et al., 2003). As a consequence, the influence of these funds is also growing. The shareholders can have a bigger influence on the firms and demand higher CSR investments. Not only wanting to invest socially responsible but wanting the company they invest in to become more socially responsible itself.

Nowadays, SRI is not only done by SRI funds, however, it is also a common investment strategy for institutions and private investors. More and more countries and large institutional investors adopted

SRI policies, which led to divestments of controversial companies and investments in ethical companies, for example, the Norwegian sovereign wealth fund (Vasudeva, 2013).

2.3.2 Measuring SRI

As seen in previous research regarding SRI, SRI can be measured in different ways. Regarding funds, funds can screen their investments in two ways, positive screens, and negative screens. A positive screen looks at firms that are performing well in environmental, social, and ethical categories. Funds choose these companies to invest in. For negative screens, funds exclude stocks that perform poorly in these categories. A good example is sin stocks, sin stocks are firms in controversial industries such as alcohol, tobacco, and weaponry.

How does a fund then qualify as an SRI fund? Rating agencies or data sources such as Morningstar give out ratings regarding the sustainability of a fund. Morningstar gives a rating from 1 to 5. Where a 1 is the bottom 10% and a score of 5 is the best performing 10% regarding SRI. However, these ratings are controversial and not always in line with other rating agencies, leading to heterogeneity among SRI scores (Durán-Santomil et al.,2019). These scores are calculated by looking at the portfolios of the funds and scoring them on their screenings and the percentage invested in SR stocks. In research, the Morningstar rating is a popular choice.

For stocks, several ways to measure their social responsibility have been implemented. Firms were screened, as previously explained, in two ways. In several categories, these firms are scored based on negative and positive screens. KLD is a rating agency that measures the social responsibility of a company. KLD used multiple criteria to calculate the scores. Qualitative and exclusionary criteria. For qualitative criteria, KLD looks for positive actions in criteria such as human rights, diversity, and environment. The exclusionary rating checks how many controversial businesses the company is involved in (Kempf & Osthoff, 2007). In earlier research, firms were categorized based on the Innovest Strategic Value Advisors' eco-efficiency scores, where firms are scored on the amount of waste generated against their total production output (Derwall et al., 2005).

Nowadays, ESG scores are popular. ESG scores are somewhat similar to the KLD scores. ESG stands for Environmental, Social, and Governmental. Companies are screened in all these categories and given a rating between 0-100, where 100 is the highest ESG score. ESG scores are given by multiple rating agencies which can be contradicting. Therefore, ESG scores can be controversial and not concluded. In research, this leads to different results when different rating agencies are used for ESG scores.

3. Data & Methods

3.1 Data on sustainability scores

The main variable of interest in this study is the sustainability score of a company or mutual fund. Sustainability scores are measured in several ways and different methods are used to construct these scores.

The first hypothesis requires mutual fund data. Morningstar provides extensive data on mutual funds and their sustainability performance. I have retrieved the sustainability scores for each fund for October 2020. The variable used to indicate the sustainability score is Morningstar Sustainability Rating (MSR), which rates each fund from 1 to 5, where 1 means a portfolio with the highest ESG risk and 5 means a portfolio with the least ESG risk (Morningstar, 2021). This score is computed from a historical and current sovereign and corporate sustainability score, as shown in figure A1 of the Appendix. These scores are then multiplied by the percent contribution of sovereign or corporate assets of their portfolio. Since I am more interested in equity than fixed-income and in corporate stocks and bonds rather than sovereign bonds I have only selected funds with a portfolio that have 100% corporate exposure. These funds' corresponding MSR only depends on their Corporate Sustainability Score. This score is computed by Sustainalytics, which gives each holding a Risk Rating based on their unmanaged ESK risk. All portfolio holdings are analyzed, and Corporate Sustainability Score (CSS) is produced. This CSS is then transformed into an MSR together with the historical Corporate Sustainability Scores.

Besides the MSR, I will use the Corporate Sustainability Percent Rank in the Global category. This variable ranks all funds based on the sustainability of their corporate holdings in their portfolio. This variable has values from 1 to 100, where 1 is the top 1% of companies regarding corporate sustainability.

The sample consists of only open-ended funds labeled as equity funds. I required a minimum asset allocation to equity of 80% as an extra measure. Furthermore, the minimum fund size is 10,000,000 USD and the fund must be actively managed and based in Europe, Asia, or North America. Lastly, only funds with an MSR were included in the sample. This resulted in a sample of 12,832 funds.

For the second and third hypotheses, I have used the Refinitiv Asset 4 database to retrieve the Thomson Reuters ESG scores. The Thomson Reuters ESG scores are designed to measure a company's relative ESG performance, commitment, and effectiveness. Refinitiv covers 10 main themes and minimalizes size biases. The scores are from 0-100 where 0 is the poorest relative ESG performance score and 100 is the highest relative ESG performance. The performance is relative to the company's

country and sector (Refinitiv, 2022). The ESG scores of the entire universe from 2002-2020 have been retrieved for the second hypothesis. For the third hypothesis, the ESG scores of 2021 are used. The data covers more than 10,000 companies across 76 countries. This dataset includes SIC codes, industries, countries, and stock exchange(s) variables. After removing missing ESG scores and dropping financial companies with a SIC code of between 6000-6999, the base dataset contains 54,453 observations across 7,277 firms.

3.2 Data on financial performance

The data for the performance of the mutual funds have also been retrieved from Morningstar. The fund's annual return for 2021 has been used as one of the performance measures. Besides the fund's annual return, the net expense ratios are downloaded to create the investors' annual return. Since investors pay an annual fee to the funds, which is the net expense ratio, the annual investors' return is the annual funds' return minus the net expense ratio. This variable is called the Annual investor return. Furthermore, multiple control variables are used to determine the relative performances. First, regarding the region of the fund, the three possible regions are Europe, North America, and Asia. There might be differences in financial performance caused by the region rather than the sustainability aspect of the fund. Moreover, the Morningstar Benchmark is considered to control different fund types. Besides the Morningstar Benchmark, Morningstar also assigns an equity style to each fund. This is a 3x3 box based on the size and book-to-market ratios of the underlying stocks of the fund. The three sizes are 'Small', 'Mid', and 'Large'. The book-to-market ratios are divided into 'Growth', 'Blend', and 'Value'. Lastly, the fund size and turnover ratio control differences in size and the level of active management. overcome fund size bias and asset allocation bias.

For the second hypothesis, data containing monthly stock return indices have been used. I have used the return index instead of the prices of companies to account for dividends or interests that could be reinvested. This data is retrieved from Refinitiv Datastream. The ESG firm-level data is combined with this dataset to create a monthly panel dataset including ESG scores and monthly returns. After removing missing ESG and return data, the dataset contains 632,467 observations across 7,112 firms. The risk-free rate has been used to create Excess returns; this risk-free rate is collected from *Kenneth R. French's library* (2022). The control variables used are based on the three-factor model of Fama & French (1993) and the four-factor model of Carhartt (1997). These control variables are small minus big (SMB), which controls for the phenomenon that small firms tend to outperform large firms over time; high minus low (HML), which accounts for the outperformance of value stocks; Excess market return (Rm-Rf), which is the average market return; at last, momentum (MOM), this factor is based on

the occurrence that stocks that had a positive return in the past will remain positive and past stocks that had a negative return in the past will remain negative.

The above has been replicated for my event study in the third hypothesis. however, there are differences concerning the interval and the period. Since I will execute an event study, I collected data around the earnings announcement dates, which I have chosen for the third quarter of 2021, due to data availability and it is the most recent data available. Therefore, the earnings announcement dates were within the first two months after the second quarter. For completeness, return data of at least 250 trading days before the first announcement date and 20 days after the last announcement date have been retrieved. The variables concerning the period are announcement dates, fiscal period, and fiscal year. To determine the earnings surprises, SUE is used, which is Standardized Unexpected Earnings, which is based on the Earnings per Share (EPS). This data is retrieved from I/B/E/S, via WRDS. The stock price data for 2021 is gathered via Refinitiv Eikon Datastream. After retrieving this data for all companies in the ESG database, I narrowed it down to companies based in the United States. This is done for two reasons: I expect more analysts to estimate the earnings of companies based in the United States. This prevents the situation where one analyst's estimation can heavily impact the SUE.

3.3 Methods

This section will provide the method for how the hypotheses will be tested. The decision behind the methods of analysis and the methods will be explained. Three methods will be described, one for each hypothesis.

3.3.1 Method hypothesis 1

To find the difference in returns between conventional and ESG funds, I need to determine what ESG funds are. Firstly, I will perform a linear regression using OLS with the Sustainability rating as the independent variable. The Sustainability rating will be a categorical variable with 5 values. This means that funds will not be distinguished as ESG and conventional but can take 5 values ranked from low to high sustainability scores. The regression will determine whether a higher Sustainability rating leads to a higher return. The regression equation is then computed as follows:

(1) $R_{f,t} = \alpha + \beta * Low_{t-2} + Below\ Average_{t-2} + Average_{t-2} + Above\ Average_{t-2} + {X'}_{t-2}\beta_2 + \varepsilon$

(2) $R_{f,t} = \alpha + \beta * Sustainability\ Rating_{t-2} + {X'}_{t-2}\beta_2 + \varepsilon$

Where $R_{f,t}$ is the return of fund f at time t and ${X'}_{t-2}$ is a vector for certain control variables. $Sustainability\ Rating_{t-2}$ is the Morningstar Sustainability rating at time $t-2$. The sustainability rating can have the values 1-5. Where 5 is the best sustainability rating and 1 is the worst. The

sustainability rating and other control variables are lagged for two months. The Sustainability variable will first be used as a dummy variable, where all 5 ratings will be compared to each other in the regression. This is shown in equation 1, where the different ratings are specified except for the 'High' ranking, this means that the coefficients of the other four rankings are relative to 'High'. I expect the coefficients of these rankings to be positive, with an emphasis on the lowest rankings. The variable can also be used as a continuous variable with values 1 to 5 which represent ascending sustainability ratings (Equation 2). The hypothesis states that conventional funds outperform ESG funds. Therefore, I expect the coefficient β to be significantly negative for equation 2, meaning that the higher the Sustainability rating, the lower the returns.

Furthermore, I will also use the Corporate Sustainability Percent Rank to determine the performance of sustainable and conventional funds. The corresponding regression equation is

(3) $R_{f,t} = \alpha + \beta * Corp.Sust.\%\, Rank_{t-2} + X'_{t-2}\beta_2 + \varepsilon$

The difference with the second equation is the variable $Corp.Sust.\%\, Rank_{t-2}$. The hypothesis states that conventional funds outperform ESG funds. Therefore, I expect the coefficient β to be significantly positive, meaning that the fund's returns are lower when the fund is sustainable. The percent rankings are from 1 to 100 in descending corporate sustainability scores. Therefore, I expect a fund with a ranking of 1 to be outperformed by a fund with a ranking of 100.

To account for differences in fund types, these regressions also are executed one by one for funds with exclusively the same Morningstar benchmark. This way, only similar funds will be compared with each other, which eliminates most biases or other sources of differences in returns besides sustainability.

3.3.2 Method Hypothesis 2

The second hypothesis states that an ESG factor can help to explain abnormal returns in a (multi)-factor model. I need to create an ESG factor to understand whether this is the case. The ESG factor will be created in three different ways. I will form a High ESG minus Low ESG (ESG) portfolio for all these cases. This means that the portfolios take a long position in stocks with a high ESG score and a short position in stocks with a low ESG score.

These portfolios are then assessed by three different factor models. First, I will test the portfolios in the traditional CAPM model, introduced by Sharpe (1964) is as follows:

(4) $R_{it} - R_{ft} = a_{it} + \beta_1 (Rm - Rf)_t + \varepsilon_{it}$

Where $R_{it} - R_{ft}$ is the excess return for stock i at time t, a_{it} is the abnormal return for stock i at time t, $(Rm - Rf)_t$ is the market return minus the risk-free rate at time t, and β is the factor coefficient. As one can see from the equation, the CAPM assumes that the return of an asset solely depends on the market return and its factor coefficient. Fama & French (1993) updated this model to include new factors besides the market return. Size and Value were found to explain stock returns. This is the Fama & French (1993) three-factor (FF3) model:

(5) $R_{it} - R_{ft} = a_{it} + \beta_1(Rm - Rf)_t + \beta_2 * SMB_t + \beta_3 * HML_t + \varepsilon_{it}$

Where the added factors are SMB_t, which denotes the size premium and HML_t is the value premium. In 1997, Carhart added a factor to the model, which is the momentum factor.

(6) $R_{it} - R_{ft} = a_{it} + \beta_1(Rm - Rf)_t + \beta_2 * SMB_t + \beta_3 * HML_t + MOM_t + \varepsilon_{it}$

Where MOM_t is the momentum factor winners minus losers. These three models will be used to assess the performance of the ESG portfolios. The equal-weighted ESG portfolios are formed by dividing the sample into deciles, quintiles, and tertiles, based on their ESG scores. This will result in ten, five, and three subsets respectively. I will first compare the highest and lowest quantiles to analyze whether ESG scores predict stock returns. After that, the portfolios will be formed taking a long position in the highest quantile and a short position in the lowest quantile. These are the ESG portfolios. The return of these portfolios will then be regressed on the CAPM, FF3, and Carhart's (1997) four-factor model. The regression equation for all regression is as follows:

(7) $R_{ESG,t} = \alpha + X'_t\beta + \varepsilon$

Where $R_{ESG,t}$ is the excess return of the ESG portfolios and X'_t is a vector for other factors. The results will yield an alpha (α), that if significantly different from 0, that explains stock returns.

3.3.3 Method Hypothesis 3

The third hypothesis states that SR investors are less sensitive to earnings surprises than conventional investors. To test this hypothesis, I will conduct an event study examining US stocks' stock prices with varying ESG scores after an earnings announcement. For most companies, analysts forecast their earnings after each quarter. The actual earnings sometimes differ from the forecasted earnings, which is the earnings surprise. For this study, the EPS will be used as a measure of performance and earnings. A widely used measure for earnings surprises is the Standardized unexpected earnings (SUE). This has been introduced by Latane & Jones (1977). The SUE is the unexpected earnings deflated by the standard error of the forecasted earnings. The formula for SUE is shown below:

(8) $SUE = \frac{EPS_Q - fEPS_Q}{SE_Q}$

Where EPS_Q is the reported EPS for quarter Q, $fEPS_Q$ is the forecasted EPS for quarter Q and SE_Q is the standard error of the forecasted earnings for quarter Q. The SUE will be used in this event study to determine which companies had a positive or negative earnings surprise after the actual earnings announcement. The event study consists of two samples. The first sample are companies with a SUE below -1, this will be called 'Bad News'. The other sample consists of companies with a SUE higher than 1, which will be called 'Good News'.

An event study can measure the impact of a certain event. In this case, the event study can measure an earnings surprise's effect on a certain firm's stock price. An event study consists of an estimation period and an event window. The event of interest is the earnings announcement date. Commonly, the event window does not only include the event day, to account for late reactions or leaked information (MacKinlay, 1997). The estimation period is used to estimate the normal returns of a certain company before the event. The event window looks at the returns during and around the days of the event to find potential abnormal returns that have been caused by the event. The abnormal return for firm i at event date t is

(9) $AR_{i,t} = R_{i,t} - E(R_{i,t}|X_t)$

Where $AR_{i,t}$ is the abnormal return, $R_{i,t}$ is the actual return and $E(R_{i,t}|X_t)$ is the normal return. The normal return is the expected return based on the conditional X_t. This is a constant in the *constant mean return model,* where it is assumed that the mean return of a firm is constant through time. t is the event day. The expected return is de average return in the estimation period.

When using the market model, the abnormal returns for stock i are as follows:

(10) $AR_{i,t} = \alpha - R_{i,t} - \beta M_t$

M_t is the return of the market in this case, the MSCI US index' returns, and β is the market Beta. The market model will also be used to estimate abnormal returns.

After calculating the AR, CARs (Cumulative Abnormal Returns) can be constructed. CARs are the sum of all abnormal returns. These are calculated over all the days in the event window.

(11) $CAR_T = AR_{t1} + AR_{t2} + \cdots + AR_T$

Where T is last day of the event window. Since I have multiple firms, I then create Cumulative Average Abnormal Returns (CAAR):

(12) $CAAR_t = \frac{1}{N}\sum_{i=1}^{N} CAR_{i,t}$

Where N is the number of observations. The $CAAR_t$ is the average of all $CAR_{i,t}$ in the sample. Finally, to test for significance, the Cross-Sectional test is performed to test the significance of the CAAR. The cross-sectional test tests whether CAAR is significantly different from 0. The test statistic is given by

(13) $t_{CAAR} = \sqrt{N}\frac{CAAR}{\sigma_{CAAR}}$

Where σ is the standard deviation of CAAR, which is the square root of the variance which is calculated as follows:

(14) $\sigma CAAR_t = \frac{1}{N-1}\sum_{i=1}^{N}(CAR_{i,t} - CAAR_t)^2$

The t-statistic finds whether the CAARs are significantly different from zero in the mean return and market models. When the CAAR is different from zero in the mean return model, it means that the event study has had a significant effect on the stock price based on the price trend before the event. In the market model, a CAAR different from zero means that the stock price trend is significantly different from the market index price trend.

However, to test the differences between the CAARs of high and low ESG firms. An unpaired t-test is executed. The formula to calculate the test statistic is as follows:

$$t = \frac{CAAR_{HighESG} - CAAR_{LowESG}}{SE_{HighESG} + SE_{LowESG}}$$

Where SE is the standard error. The CAARs are formed by averaging all CARs of the High and low ESG groups.

Previous research papers have been consulted to determine the appropriate event window and estimation period. Multiple papers have used an estimation period of 250 trading days which ends between 40 and 45 days before the event date. MacKinley (1997) has written the paper Event Studies in Economics and Finance in which general rules for event studies are explained. MacKinley uses 250 trading days for the estimation period, which does not overlap with the event window. The event window starts 20 days before the announcement date and ends 20 days after the announcement date. McWilliams & Siegel (1997) and Moorman & Lehman (2004) recommend using 255 and 250 trading days as the estimation period ending 46 and 45 days before the event date respectively. Based on these papers, I will construct an estimation period of 250 days ending 45 before the announcement date. The event window varies between studies and depends on the type of event. Therefore, several event windows will be used. In line with MacKinley (1997), I will construct a {T-20, T+20} event window. I will also use a {T-1, t+4} event window as used by Armitage (1995), who conducts a similar study.

3.4 Descriptive statistics

Table 1 shows the descriptive statistics of the mutual funds for hypothesis 1. The funds in this sample have an average return of more than 15% and an average investor return of 14.5%. These returns vary a lot between firms, as the maximum return is 83.62% and the lowest return is -32.97%. The sample consists of funds in all five categories of the Morningstar Sustainability Ranking, where the average MSR is the most represented with 33.5%. The below-average and above-average rankings make up 21.1% and 28.1% of the sample. 9.8% of the sample has the highest MSR and 7.3% has the lowest MSR. This is in line with the ranking procedure as the MSR is distributed as follows (from high to low MSR); Best 10%, next 22.5%, Next 35%, next 22.5%, worst 10%

The average fund size is 2.925 billion US dollars, and the average turnover ratio is 45%. Almost the entire sample consists of European and North American funds as only 1.5% are Asian funds. Regarding the funds' equity styles, some are more represented than others. The Large Growth equity-style funds represent 35.5% of the funds in the sample. Funds that mainly invest in large stocks account for 72.2% of the sample in total. The least represented equity styles are Small Value and Mid Value which account for 2.9% and 3.1% of the sample respectively.

Table A1 of the Appendix shows the correlation matrix of the variables of the first hypothesis. For clarity reasons, only three of the nine equity styles have been included in the table. There are no surprising results regarding the correlations. None of the variables show a correlation higher than 0.3 or lower than 0.3 except for the obvious correlations between Annual return with the Annual investor return, MSR with the corporate sustainability percent rank and the negative correlation between Europe and North America.

Table 1 Descriptive statistics of fund characteristics and returns

Variable	Obs	Mean	Std. Dev.	Min	Max
Annual return %	11,514	15.310	12.839	-32.974	83.616
Annual investor return %	8,152	14.607	13.143	-34.763	82.466
Corp. Sust. % rank	11,627	46.132	28.220	1	99
Asset Allocation Equity %	12,832	97.343	2.907	80.366	134.659
Net Expense Ratio	8,519	1.174	.547	0	4.440
Fund size in Billion $	12,832	2.791	10.739	.010	139.664
Turnover Ratio %	7,768	46.353	44.460	-105	584
Low	11,627	.073	.260	0	1
Below average	11,627	.213	.410	0	1
Average	11,627	.334	.472	0	1
Above average	11,627	.281	.449	0	1
High	11,627	.098	.298	0	1
Europe	12,832	.462	.499	0	1
North America	12,832	.524	.499	0	1
Asia	12,832	.015	.120	0	1
Small value	12,832	.029	.169	0	1
Small growth	12,832	.039	.193	0	1
Small blend	12,832	.035	.185	0	1
Mid value	12,832	.031	.174	0	1
Mid growth	12,832	.071	.256	0	1
Mid blend	12,832	.072	.259	0	1
Large value	12,832	.135	.341	0	1
Large blend	12,832	.232	.422	0	1
Large growth	12,832	.355	.479	0	1

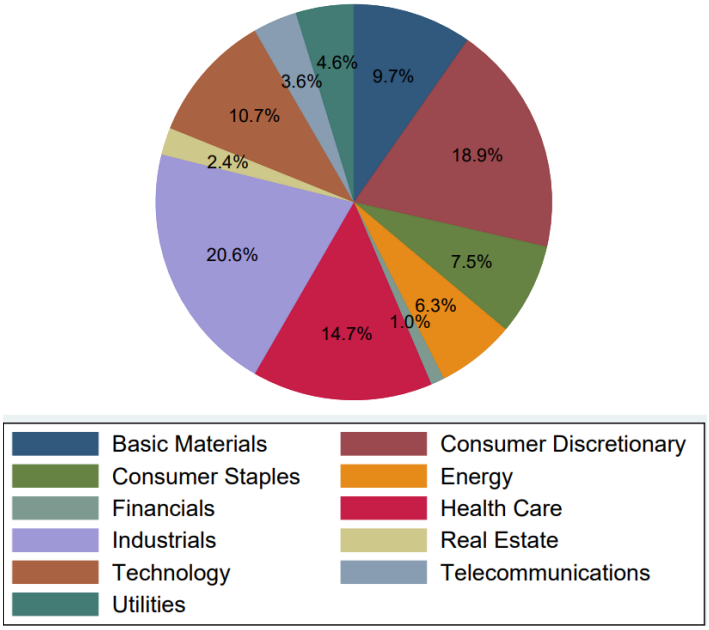
Next, I computed the average of all variables within each MSR to find the first sign of evidence for my first hypothesis, but also how the MSR is related to the other variables. Table 2 shows the average values of the variables. Some interesting findings are the differences in turnover ratio and the fund size. Higher MSR funds seem to have a lower turnover ratio and a smaller fund size. Also, Europe seems to have a relatively higher percentage of MSR funds than North America. Regarding the equity styles, High MSR funds are for 57.3% funds with a Large Growth equity style. On the contrary, Low MSR funds are for only 10.9% Large Growth funds. When looking at all the columns, the percentage of Large Growth funds seems to increase with a better MSR. Large Value and Small Value funds, on the other hand, seems to decrease with a better MSR. Lastly, Lower MSR funds seem to have a higher annual return than high MSR funds, however, these results should be interpreted cautiously.

Table 2 Average values within each ranking of the Morningstar Sustainability Ranking

Variable	Low MSR	2	3	4	High MSR
Annual return %	14.466	17.845	15.783	13.980	13.412
Annual investor return %	14.833	17.513	14.939	12.370	13.749
Corp. Sust. % Rank	93.651	76.283	48.688	21.468	7.328
Asset Alloc. Equity %	96.609	97.305	97.212	97.495	97.926
Net Expense Ratio	1.139	1.151	1.171	1.201	1.197
Fund size in Billion $	3.428	3.070	3.812	2.154	1.769
Turnover Ratio %	60.471	49.921	49.351	38.534	33.085
Europe	.400	.398	.406	.508	.498
North America	.594	.598	.585	.481	.485
Asia	.006	.003	.008	.011	.017
Small value	.068	.040	.034	.020	.005
Small growth	.089	.025	.041	.033	.048
Small blend	.052	.032	.045	.035	.024
Mid value	.060	.045	.033	.014	.030
Mid growth	.111	.048	.057	.089	.060
Mid blend	.066	.089	.075	.069	.039
Large value	.272	.217	.146	.070	.032
Large blend	.174	.251	.251	.230	.187
Large growth	.109	.254	.318	.441	.573

For the second and third hypotheses, I retrieved the Thomson Reuters ESG scores via the Refinitiv Asset 4 database. This database included the industries of the companies as well as the stock exchanges listed. Figure 1 shows the represented industries of firms in the complete dataset. The most represented industries are Industrials and Consumer Discretionary which account for 20.6% and 18.9% of the sample respectively. 14.7% of the companies in the sample are in the Health Care industry. An interesting finding is that, even after removing SIC Codes of 6000-6999, which represents the Finance and Real Estate sector, the sample still consists of 1% Financials and 2.4% Real Estate. After further investigating these sectors. The firms Real Estate are mostly Operative Builders and the firms that are in Finance according to the data are not always in line with their SIC Codes.

Figure 1 Industries represented in the sample



As for the return and ESG data, the summary statistics are shown in Panel A of table 3. The average ESG score in the sample is 41.89 and the average return is 1.3% per month, meaning that on average stock prices have increased 1.3% in this sample from the years 2002-2020. The lowest return is -100%, meaning that a stock has lost all its value. The highest return is more than 800%, that certain stock had become eight times more valuable in one month.

Panel B shows the factors of the factor models used in hypothesis 2. All the factors generated positive returns, except for the HML factor, which generated a monthly negative return of 0.2%. The high-Low ESG portfolios have a negative return of 0.5% a month on average, this is the case for the deciles and quintiles. The High-Low ESG portfolio of tertiles generates a negative return of 0.4% a month. These results should be interpreted carefully, however, this could be the first evidence of the second hypothesis.

Table 3 Descriptive Statistics Funds, Factors & High Low Portfolios

Variable	Obs	Mean	Std. Dev.	Min	Max
Panel A: Fund level					
ESG score	632,465	41.89	20.773	.14	95.19
Return	632,465	.013	.128	-1	8.408
Panel B: Factor statistics					
Mkt-Rf	228	.007	.044	-.172	.137
SMB	228	.002	.024	-.053	.060
HML	228	-.002	.027	-.139	.082
MOM	228	.001	.047	-.343	.123
Panel C: High-Low ESG portfolios					
Deciles	226	-.005	.022	-.095	.090
Quintiles	226	-.005	.019	-.086	.074
Tertiles	226	-.004	.015	-.075	.054

Regarding the event study of hypothesis 3, Table 4 shows the descriptive statistics for the entire sample and for the two samples split into 'Good news' and 'Bad news'. Panel A shows the summary statistics of the entire sample and Panel B and Panel C show the summary statistics of the 'Bad news' and 'Good news' samples respectively. The overall mean of SUE is surprisingly high, with an average of 1.640. This means that, in this sample, analysts generally underestimate the quarterly earnings of a firm. The number of firms that has a SUE higher than 1 is substantially higher than the firms with a SUE lower than -1 (1,221 to 430). The average ESG score of the good news firms is also higher than firms with bad news, with a difference of almost 8 on a scale of 100. For all samples, there are some extreme SUEs, however, I decided to leave them in, since there will not be more weight on extreme SUEs in the event study.

Table 4 Descriptive Statistics of the firms used in the event study

Variable	Obs	Mean	Std. Dev.	Min	Max
Panel A: Entire sample					
ESG score	2,493	45.244	18.501	2.060	94.43
SUE	2,493	1.640	6.479	-47.94	137.887
Panel B: Bad News					
ESG score	430	40.339	18.184	9.170	92.000
SUE	430	-4.081	5.333	-47.94	-1.007
Panel C: Good News					
ESG score	1,221	48.011	18.596	2.060	92.510
SUE	1,221	4.718	7.271	1.001	137.887

4. Results

This section shows the results for all three hypotheses. These hypotheses have been tested as explained in section 3.

4.1 Hypothesis 1

Multiple regressions have been executed to test the first hypothesis and are shown in table 5. Two dependent variables are used: Annual return and Annual Investor return. The difference between the variables is the actual return that the investors receive versus the theoretical return that the portfolio has generated throughout the year. The difference between these two measures is the net expense ratio, which is the annual fee investors pay the fund. Also, two measures of SR funds are used: the Morningstar Sustainability Rating and the Corporate Sustainability % rank. Panel A of table 5 shows the regressions with annual return as the dependent variable. First, two regressions are performed with the MSR as the independent categorical variable. "High" is omitted, therefore the coefficients of the other four categories are relative to the Highest MSR. Column 1 is a regression of the annual return on only the MSR. All four categories have a positive coefficient, meaning that the funds with the lowest four MSR categories (Low, Below average, Average, and Above average) have higher returns than funds with the highest MSR. However, only the average and below average MSR are significant on a 99% confidence level and the Low category on a 90% confidence level. When adding control variables in column 2, funds with the highest MSR are significantly outperformed by all other categories, however, only Low, Below average, and Average are significant on a 99% confidence level. Low outperforms High by 2.040%, Below average, and average by 4.456% and 2.458% respectively. Funds with the highest MSR have the lowest returns while the below-average funds have the highest returns.

Another interesting finding is the effect of the region of the funds. Europe is omitted, meaning that the coefficients of Asia and North America are relative to Europe. Asia has a coefficient of -12.923, which is significant on a 99% confidence level. This means that Asian funds are significantly outperformed by European funds. In this sample by almost 13% a year. On the other hand, North American funds significantly outperform European funds by almost 5% annually. The fund size and turnover ratio do not seem to influence the annual return.

Columns 3 and 4 show the same regressions, however, the Corporate Sustainability Percent Rank is used as the sustainability variable. When the Annual return is regressed on only the Corp. Sust. % Rank. The coefficient is 0.025 and significant on a 99% confidence level. This means that the lower the percent rank the fund has, the return increases by 0.025%. To make this even more concrete; a top 1% sustainability fund has an annual return of 14.248% (constant=14.223), and a fund in the 41st percentile

has a return of 15.248%. When adding the control variables, the coefficient of the sustainability variable does not change and is still significant on a 99% confidence level. The control variables have the same effect as in column 2. Asian funds have a significantly lower return than European funds and North American funds significantly outperform European funds.

Panel B of table 5 shows the effect of sustainability on Annual investor return. Column 1 shows that Below average and Average significantly outperform High on a 99% and 95% significance level respectively. Even more, when adding the control variables, the latter also becomes significant on a 99% significance level. An interesting finding is that the Above average MSR funds seem to outperform the High MSR funds in column 1. However, when adding the control variables in column 2, the coefficient is not significant anymore. In this sample, the annual fund return for the High MSR funds is the lowest of all the rankings. However, this effect decreases when subtracting the net expense ratios to create the Annual investor return.

Columns 3 and 4 of panel B again show the Corporate Sustainability Percent Rank as the sustainability variable. The results of these regressions are somewhat similar to columns 3 and 4 of Panel A. The coefficient remains positive and significant on a 99% confidence level. The effect of the regions is almost the same for all regressions, even more, the effect of the fund size and turnover ratio is not significant in panel B either.

Table 5 Regressions of Annual return and Investor return on MSR categories and Corp. Sust. % rank. (Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1)

Panel A.	(1) Annual Return(%)	(2) Annual Return(%)	(3) Annual Return(%)	(4) Annual Return(%)
Low	1.055*	2.040***		
	(0.589)	(0.722)		
Below average	4.433***	4.456***		
	(0.464)	(0.591)		
Average	2.371***	2.458***		
	(0.437)	(0.561)		
Above average	0.568	0.954*		
	(0.447)	(0.572)		
Corp. Sust. % rank			0.025***	0.023***
			(0.004)	(0.005)
Asia		-12.923***		-13.432***
		(2.152)		(2.519)
North America		4.888***		5.176***
		(0.446)		(0.446)
Fund size in Billion $		-0.008		-0.009
		(0.010)		(0.010)
Turnover Ratio		-0.004		-0.003
		(0.003)		(0.003)
Constant	13.412***	11.300***	14.223***	12.164***
	(0.384)	(0.633)	(0.233)	(0.462)

Panel B.	(1) Investor return(%)	(2) Investor return(%)	(3) Investor return(%)	(4) Investor return(%)
Low	1.084	0.777		
	(0.737)	(0.809)		
Below average	3.764***	3.738***		
	(0.567)	(0.667)		
Average	1.190**	1.717***		
	(0.539)	(0.625)		
Above average	-1.379**	-0.241		
	(0.555)	(0.640)		
Corp. Sust. % rank			0.039***	0.022***
			(0.005)	(0.006)
Asia		-13.526***		-13.863***
		(2.222)		(2.226)
North America		4.996***		5.240***
		(0.502)		(0.502)
Fund size in Billion $		-0.013		-0.015
		(0.011)		(0.011)
Turnover Ratio		-0.003		-0.001
		(0.004)		(0.004)
Constant	13.749***	12.895***	10.794***	10.817***
	(0.479)	(0.284)	(0.721)	(0.519)

Furthermore, the results of regression 2 with MSR as a continuous variable are shown in Table 6. Columns 1 and 2 show the regressions of the Annual return on only MSR and MSR with several control variables respectively. Columns 3 and 4 show the same regressions with Investor return instead of Annual return. The results are in line with the previously discussed results of table x, In all four columns, MSR has a significant negative coefficient. The magnitude of this coefficient is around 1% for all four columns. Column 4 for example, shows a significant coefficient of -0.950 for MSR. Meaning that a category of MSR higher leads to a 1% lower annual investor return. However, Panel B of table x has shown that this decrease is not linear, since the lowest MSR seems to be outperformed by the second and third lowest MSR category.

Regarding the control variables, this regression is completely in line with the previous regressions. North America outperforms Europe, meanwhile, Asia underperforms. The turnover ratio and the Fund size do not seem to affect the annual returns.

Table 6 Annual return and Investor return on Morningstar Sustainability Rating with control variables

	(1)	(2)	(3)	(4)
VARIABLES	Annual Return(%)	Annual Return(%)	Investor return(%)	Investor return(%)
MSR	-1.010***	-1.034***	-1.303***	-0.950***
	(0.112)	(0.138)	(0.137)	(0.155)
Asia		-12.879***		-13.272***
		(2.154)		(2.224)
North America		5.075***		5.248***
		(0.445)		(0.501)
Fund size in Billion $		0.009		0.015
		(0.010)		(0.011)
Turnover Ratio		-0.004		-0.003
		(0.003)		(0.004)
Constant	18.539***	16.569***	18.744***	14.857***
	(0.369)	(0.633)	(0.449)	(0.697)
Observations	11,159	7,152	8,015	6,054
R-squared	0.007	0.038	0.011	0.038

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

4.2 Hypothesis 2

Hypothesis 2 states that an ESG factor explains differences in cross-sectional stock returns. To test this hypothesis, High minus low ESG firm portfolios have been formed. The return of these portfolios is then regressed on several other factors to test the significance of the alpha of the return of this portfolio. Table 7 shows the results for the second hypothesis. Column 1 shows the coefficients of the regression of the returns of the High-Low deciles portfolio on the CAPM model. The alpha has a significant coefficient of -0.005% on a 99% confidence level, meaning that the Low ESG decile outperforms the High ESG decile by 0.5% per month, which cannot be explained by the CAPM model. When adding the SMB and HML factors in column 2 and the momentum factor in column 3, the alpha remains -0.005. The alpha is significant in all three models on a 99% confidence level. The FF3 and Carhartt (1997) model do not explain the underperformance of a High ESG minus Low ESG portfolio sorted in deciles. Columns 4-6 show regressions of the High-Low quintiles portfolio and columns 7-9 show the regressions of the High-Low tertiles portfolio. For all types of quantiles, the alpha shows a significant negative return for the High-Low portfolios and a coefficient that does not differ more than 0.1 percentage point per model. The multi-factor models FF3 and Carhartt (1997) cannot explain the underperformance.

The market does not have a significant effect on the portfolio's return, meaning that the portfolio does not load on the market factor. The portfolios, therefore, are well diversified regarding the market risk. The other factors do not seem to have a significant effect on the returns of the High-Low portfolios. Only the SMB factor seems to negatively affect the High-Low quintile portfolio in the Carhartt (1997) four-factor model.

When portfolio sorting funds based on their ESG scores and creating a portfolio that goes long in high ESG stocks and short in low ESG stocks, it generates significant negative returns. This is the case for all three types of portfolio sorts: deciles, quintiles and tertiles. Several multi-factor models cannot explain the significant negative alphas these portfolios have generated.

Table 7 ESG Portfolio Regressions on (multi-)factor models

Variable	(1) deciles High-Low	(2) deciles High-Low	(3) deciles High-Low	(4) quintiles High-Low	(5) quintiles High-Low	(6) quintiles High-Low	(7) tertiles High-Low	(8) tertiles High-Low	(9) tertiles High-Low
Mkt-Rf	-0.023	-0.009	0.012	-0.014	-0.003	0.010	0.002	0.003	0.006
	(0.032)	(0.035)	(0.038)	(0.028)	(0.030)	(0.033)	(0.023)	(0.024)	(0.026)
SMB		-0.092	-0.095		-0.089	-0.091*		-0.045	-0.046
		(0.063)	(0.063)		(0.054)	(0.054)		(0.044)	(0.044)
HML		0.028	0.052		0.042	0.058		0.053	0.056
		(0.056)	(0.058)		(0.048)	(0.050)		(0.039)	(0.041)
MOM			0.052			0.034			0.007
			(0.036)			(0.031)			(0.025)
α	-0.005***	-0.005***	-0.005***	-0.005***	-0.004***	-0.005***	-0.004***	-0.004***	-0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Obs	226	226	226	226	226	226	226	226	226
R^2	0.002	0.012	0.021	0.001	0.015	0.021	0.000	0.012	0.012

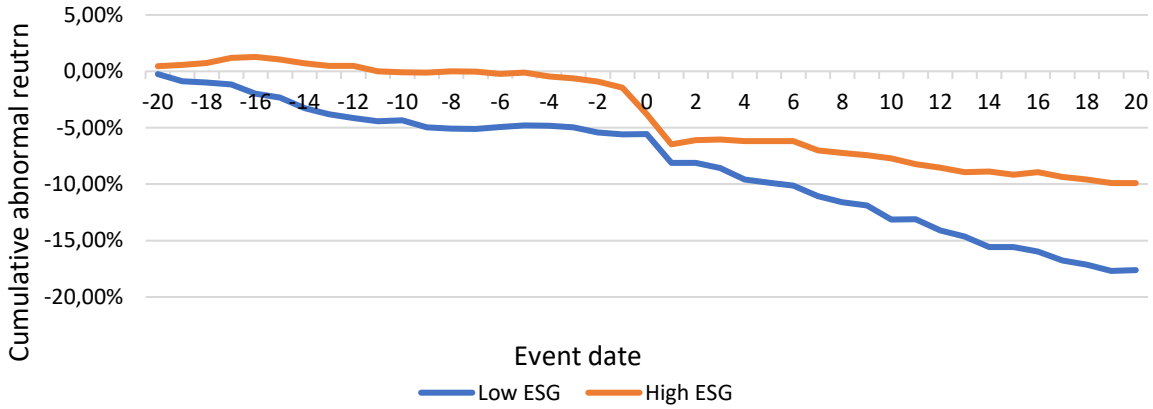
Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

4.3 Hypothesis 3

Hypothesis 3 states that SR investors are less sensitive to earnings surprises, which shows itself in the way that high ESG stocks' returns do not fall as much as low ESG stocks' returns after a negative earnings surprise. First, the CARs are calculated, for the High ESG and low ESG groups, for multiple event windows and for both negative and positive earnings surprises. CAARs are calculated by averaging the CARs within each group. Graph 1 shows the effect of negative SUEs on the returns with the CAARs market return model for an event window of {T-20, T+20}. Over the 41-day event period, it can be seen that the Low ESG stocks have a lower CAAR than high ESG stocks. This is for every day in the event period. However, around the day of the event, at t=0, the shock seems to be larger for High ESG stocks. Over the entire 41-day period, the CAARs are -9.91% and -17.63% for High ESG and Low ESG respectively, the CAARs are shown in table A2 of the Appendix. The t-statistics of High and Low ESG are -6.97 and -8.69 meaning that both show significant negative CAARs on a 1% level, their corresponding p-values are shown in table 8. The t-test with

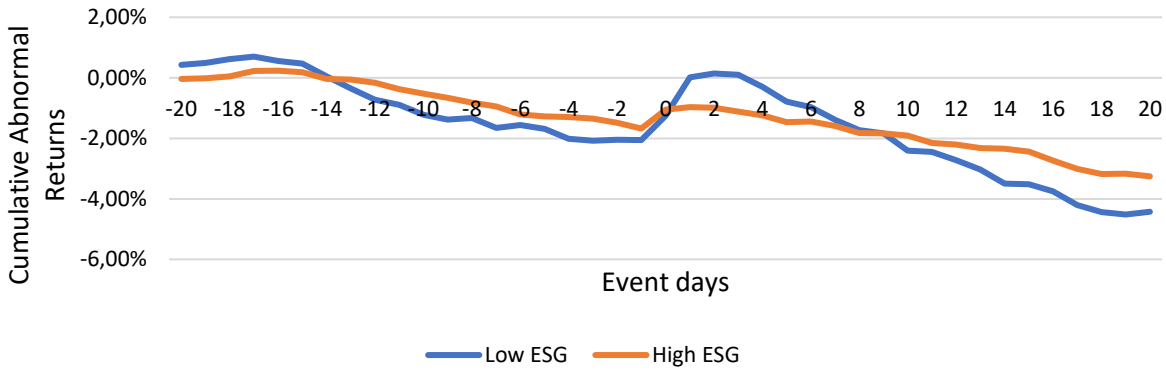
$H0 = |CAAR_{HighESG}| - |CAAR_{LowESG}| = 0$ is rejected on a 1% level with a P-value of 0.001. Therefore, for a {T-20, T+20} event period, the prices of High ESG stocks are less sensitive to bad news. This is in line with the hypothesis that SR investors are less sensitive to negative earnings surprises.

Graph 1 {T-20, T+20} CAAR Market Model for negative SUEs



For positive SUEs, this event study is repeated and shown in Graph 2. Both of the groups show an increase in returns around the event day, which stagnates before it declines again. Table A3 of the Appendix shows that the CAAR for Low ESG is -4.43% and for High ESG is -3.26%. The t statistics for Low ESG and High ESG are -4.78 and -4.93. Meaning both are significantly different from zero. This is surprising since these stocks all had a positive SUE. A trend that can be picked up from Graphs 1 and 2 is that all the abnormal returns seem to be negative 5 days after the event day. An external event in the stock market might have caused this decline. Around the event date, the increase in abnormal returns is higher for the Low ESG stocks than for the High ESG stocks. This might indicate that low ESG stocks react more heavily to the good news at the moment of the announcement.

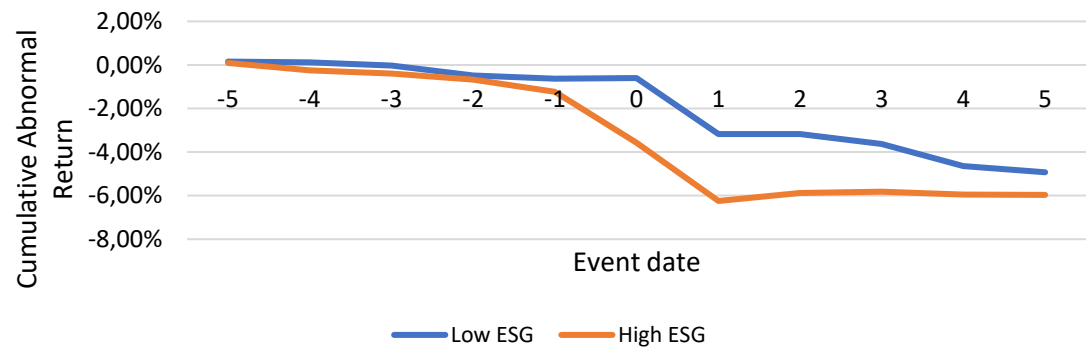
Graph 2 {T-20, T+20} CAAR Market Model for positive SUEs



The T-test to test the difference between high and low ESG stocks has a P-value of 0.304 thus there is no significant difference between high and low ESG firms regarding the effect of good news in this 41-day event window. However, the event period of this event study starts 20 days before the event day and ends 20 days after the event day. The bigger the event window, the higher the probability that other factors have caused this difference, as mentioned in the Methodology. For this reason, this event study will be repeated with shorter event windows of {T-5, T+5} and {T-3, T+3}.

Starting with the {T-5, T+5} event window, an immediate difference with the {T-20, T+20} can be noticed for bad news events in graph 3 as the high ESG stocks show a lower CAR then low ESG stocks As the difference is more than one percentage point. The immediate effect for high ESG stocks are of a larger magnitude than for low ESG stocks, which is surprising and not in line with the hypotheses that SR investors are less sensitive to bad news. However, the difference in CAR between the high and low ESG firms is not significant, as the T-test cannot reject $H0 = |CAAR_{HighESG}| - |CAAR_{LowESG}| = 0$. With a p-value of 0.52. On the other hand, both types of firms have significant CARs on a 1% level as their corresponding T-statistics are -3.61 for low ESG and -6.13 for high ESG. Therefore, investors of both firms react negatively to bad news, in an {T-5, T+5} event window.

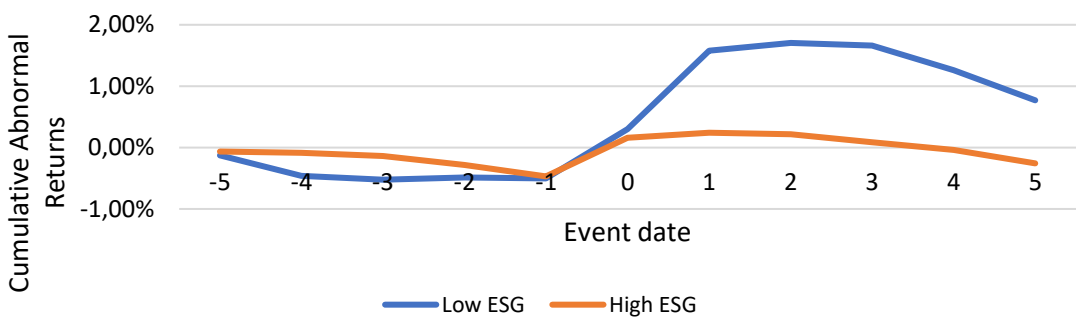
Graph 3 [T-5, T+5} CAAR Market Model for negative SUEs



CARs for the good news events are shown in graph 4. In contrast to the negative earnings surprises, High ESG firms' stock prices seem to react less to the event than those low ESG firms. Especially the day before and after the event day, the CAR of low ESG stocks increases. However, neither CARs are significant, as the T statistic for low ESG stocks is 1.28 and –0.61 for high ESG stocks. Not surprisingly, the difference in CARs is not significant either, with a P-value of 0.16.

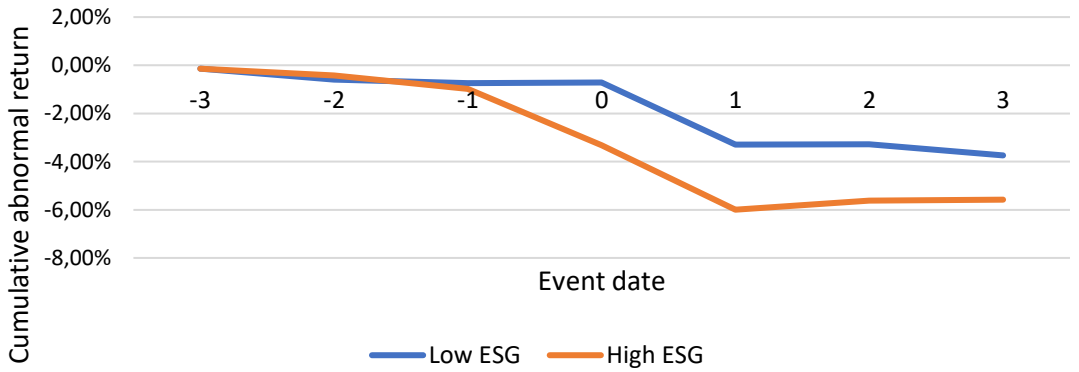
The 11-day event window has shown no differences in reaction to bad and good news between high and low ESG stocks. In general, investors are reacting more to bad news than good news. Even more, investors are only significantly reacting to bad news.

Graph 4 {T-5, T+5} CAAR Market Model for positive SUEs



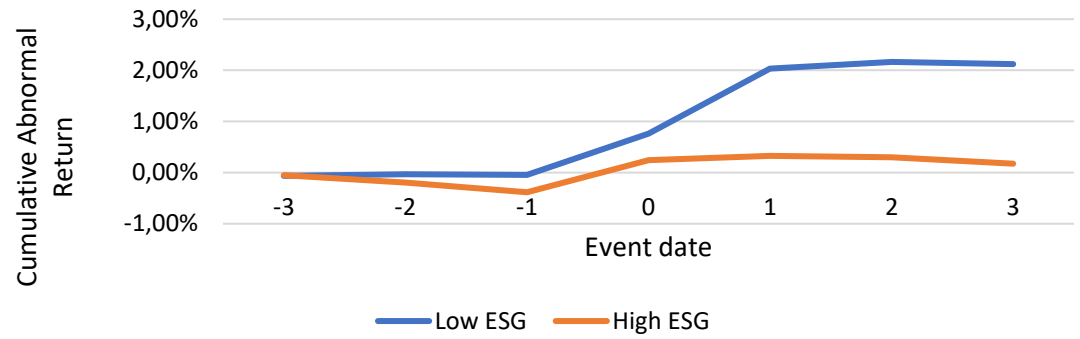
Lastly, I executed the event study with an event window of 7 days (T-3, T+3). Again, high ESG stock prices react more heavily to bad news than low ESG stock prices in this sample, which is shown in graph 5. Table A3 shows High ESG stocks have a CAR of -5.58% whereas low ESG stocks have a CAR of -3.74%. This difference is not significant, however, as the p-value is 0.21. The stock prices of both types of firms do react significantly to the bad news on a 1% level, as the t statistic of low ESG firms and high ESG firms are -3.23 and -4.84 respectively.

Graph 5 [T-3, T+3} CAAR Market Model for negative SUEs



The investors' reactions to positive SUEs in an event window of {T-3, T+3} are shown in graph 6. For high ESG stocks, this reaction is 0.17% at day T+3 with a t statistic of 0.44, as shown in table A2. Hence, there is no significant reaction of investors in high ESG firms to good news. On the contrary, low ESG stock prices have a CAR of 2.12% at day T+3, which is significant on a 1% level (T statistic = 3.95). Furthermore, the difference between the CARs is also significant on a 1% level with a P-value of 0.003. low ESG stock prices react more heavily to good news than high ESG stock prices. Even more, high ESG stock prices do not react significantly to good news.

Graph 6 [T-3, T+3} CAAR Market Model for positive SUEs



Within 1 day around the announcement date, both types of stocks show a positive reaction to good news and a negative reaction to bad news in this sample. However, for the chosen event windows, only the negative reactions to bad news are significant on a 1% level. Prices of low and high ESG stocks did not react differently to bad news, with an exception of the 41-day event window.

High ESG stock prices did not significantly react to good news in the chosen event windows, except for the 41- day event window. However, this was a negative reaction, most likely caused by external events, this is also the case for low ESG stock prices. Low ESG stock prices only reacted positively to good news in the 7-day event window. In this 7-day event window, low ESG stock prices reacted significantly more heavily to good news than high ESG stock prices.

Table 8 P-values for two-sided unpaired t-tests

	T-test H0	7-day CAR	11-day CAR	41-day CAR
Bad News	low ESG = 0	0.002***	0.000***	0.000***
	low ESG = high ESG	0.211	0.520	0.002***
	high ESG = 0	0.000***	0.000***	0.000***
Good News	low ESG = 0	0.000***	0.201	0.000***
	low ESG = high ESG	0.004***	0.163	0.304
	high ESG = 0	0.662	0.544	0.000***

5. Discussion

The main interests of this research are the behavior of Socially Responsible Investors and whether their investment preferences affect the return on their investments. The research question of this thesis is whether Socially Responsible Investing generates the same returns as conventional investing. Besides, I am also interested in whether Socially Responsible investors are as focused on financial performance as conventional investors. To answer the research question, I have stated three hypotheses.

The first hypothesis states that the performance of SR funds does not significantly differ from conventional funds. I do not find evidence for this hypothesis. Even more, I reject this hypothesis based on my findings. I find that a higher MSR significantly leads to a lower annual return in 2021 on a 1% significance level. This holds for both the funds' annual return as well as for the investors' annual return. This effect is not linear, as the least sustainable funds do not significantly outperform the most sustainable funds. Average and below-average sustainable funds are the main sources of outperformance. Even after controlling for the region of the funds, their turnover ratio, and fund size, this effect still holds. A better Corporate Sustainability Percent rank also leads to a lower annual return on a 1% significance level, this holds after incorporating the control variables and for the funds' annual return as well as for the investors' annual return. Therefore, SR funds are significantly outperformed by conventional funds, and hypothesis 1 is rejected. The found effect is in line with the findings of Cortez et al. (2012) and Duran-Santomil (2019). However, according to Duran-Santomil (2019), there is a lack of clearly defined criteria for SR funds. I have used Morningstar's methodology, which is just one angle of view. For that reason, I cannot confidently say that conventional funds outperform SR funds for all definitions of Socially Responsible funds and methodologies of measurement. Moreover, the effect is found for 2021, this limits the ability to confirm that SR funds consistently underperform conventional funds over a longer period. Lastly, The annual return of the fund is measured, whereas the sustainability of the firm can change throughout that year. This study does not control for possible changes in the sustainability profile of the fund.

I find evidence for the second hypothesis that states that an ESG factor helps explain abnormal returns in a (multi-)factor model. For all portfolio sorts, I find evidence that a High ESG minus Low ESG portfolio generates a significant negative alpha on a 1% level. This holds for all three-factor models used: CAPM (Sharpe 1964), the Fama & French (1993) three-factor model, and the Carhartt (1997) four-factor model. Besides finding that an ESG factor helps explain abnormal returns. I find that an ESG factor generates negative abnormal returns. A portfolio that goes long in high ESG stocks and short in low ESG stocks generates significant negative returns. This is in line with the findings of Belghitar et al.

(2014) and Pedersen et al. (2021) model. A possible explanation could be that ESG stocks are over-priced. The stocks are at a premium because of their green label. Cornell (2021) states a somewhat similar explanation, who says that ESG stocks are overpriced due to their popularity. I have used the ESG ratings of Thomson Reuters, again, this is just one angle of view. The ESG factor is existing in the previously mentioned models, however, over the years, multiple new well-established factors have emerged. These factors could help to explain the abnormal returns of the High ESG – Low ESG portfolio.

Lastly, the third hypothesis was tested. This hypothesis is split into two hypotheses 3a and 3b. Hypothesis 3a states that SR investors are less sensitive to a negative earnings surprise than conventional investors. I do not find evidence for this hypothesis. I only find evidence for this hypothesis through an event study using a large event window. This event window is too large to withdraw any conclusions from since the difference in stock returns is already apparent before the event. Therefore, I expect that the price changes are from an external event. Other event windows show no significant difference between stock price reactions of conventional and ESG stocks. I do however find evidence that there is a significant stock price drop after a negative earnings surprise for both conventional and ESG stocks. Concluding, hypothesis 3a is rejected. An explanation for this could be that investors in ESG stocks are not 100% SR investors, investing in an ESG stock might only have financial motives, therefore these investors would be just as sensitive to earning announcements as investors in conventional stocks. I also do not find evidence for hypothesis 3b, the sensitivity to positive earnings surprises is not higher for SR investors. Even more, for a {T-3, T+3} event window, prices of conventional stocks are more sensitive than ESG stock prices. This is surprising since a positive earnings surprise of conventional stocks would attract only conventional investors, whereas a positive earnings surprise of ESG stocks would attract both ESG investors and conventional investors, making the price increase more heavily. I do not find an explanation for this observed phenomenon. This is not in line with the existing literature on the preferences of SR investors. Renneboog et al. (2011) found that investors in SR funds were less likely to withdraw money after negative past returns. Possibly, investors of SR funds decide more actively to invest in socially responsible assets than investors in high ESG stocks. Investors in ESG stocks might not invest in it due to their socially responsible profile, but just for positive expectations about future returns. Moreover, the SUE cutoff led to a sample with much more good news cases than bad news cases. Investors seemed to react more to the bad news than the good news. This is in line with the Prospect theory that people are more sensitive to losses than to gains (Kahneman & Tversky, 1979). An equally as important note for hypothesis 1, is that this event study only studies earnings surprises of firms in the third quarter of 2021. Therefore, the results can be affected by time-varying factors or events that took place around the period of interest.

A matter that should be critically discussed is the ESG rating of firms. As mentioned, I have used one ESG rating agency to retrieve the ESG scores of the firms in this sample. Recent studies have shown that rating agencies do not agree with each other on what ESG score to give to a firm. Christensen et al. (2022) find that firms do not have the same ESG scores across the rating agencies. More differences in the ESG scores leads to more price volatility. Berg et al. (2019) find that the ESG scores of six different rating agencies, including the ratings used in this paper (Refiniv), are not in line with each other. The correlation between the scores ranges from 0.38 to 0.71. This leads to uncertainty for SR investors and the market. For this research, this decreases the scientific and social relevance of the results, as it cannot be confidently claimed that Refinitiv ESG scores represent the actual sustainability of the firms correctly. Therefore, the finding that SR funds generate lower returns becomes weaker. Furthermore, it has been found that ESG scores do not predict future socially responsible behavior (Yang, 2022). According to Bams & van der Kroft (2022), SR investors invest in high ESG firms instead of sustainable firms due to inflated ESG ratings. A high ESG portfolio even underperforms in terms of sustainability compared to the market. This is a finding that could further decrease the findings of my paper regarding the findings on the performance of SR firms. A high ESG portfolio, according to the paper is less sustainable than the market. I found that a high ESG – Low ESG portfolio generated negative returns. In a future study, where actual sustainability is used, a long short portfolio in sustainable firms might actually generate positive returns.

6. Conclusion

In conclusion, to answer the main research question, Socially Responsible Investing does not generate as much return as conventional investing based on the ESG ratings of Refinitiv. Both open-ended mutual funds and firms on the stock market do seem to be less profitable when their Morningstar Sustainability Rating and/or ESG score is higher. An OLS regression showed that a higher Morningstar Sustainability rating led to a lower annual return for 2021 in the cross-section. An ESG factor generates significant negative abnormal returns However, according to Bollen (2007), SR investors should care less about financial performance and retrieve utility from sustainable performances. Nevertheless, I do not find evidence that Socially Responsible investors are less sensitive to financial performances in an event study to find the stock price reactions on earnings surprises in quarterly announcements for the third quarter of 2021 for high and low ESG stocks.

The study on mutual funds and the event study on earnings surprises are done for 2021 and the third quarter of 2021. A future study could improve the significance by using panel data instead of only cross-sectional data. Furthermore, the sustainability ratings MSR and ESG scores are ambiguous regarding the actual sustainability of the funds and firms. It is interesting to perform this study by measuring the correct sustainability of the firm, instead of potentially overblown or inflated ratings. Where the ESG factor would be replaced by a sustainable factor. This was the first study to determine the utility function of SR investors by looking at the effect of SUEs on the stock prices of high and low-ESG firms. These results were surprising as high ESG stocks reacted less to a positive earnings surprise. It would be interesting to see a future study investigating this matter and finding possible explanations for my findings.

The findings in this paper have both practical and scientific implications. Practical implications are that firms with a high ESG score tend to generate lower returns. This is an interesting finding for all investors that can take this into account. Based on the literature that disproves that ESG scores consistently represent sustainability, I will not state that sustainability results in lower returns. Investors can take a favorable MSR into account when looking for the funds with the highest annual return, as this seems to significantly affect annual returns. A scientific implication is an ESG factor exists based on several multifactor models and should be taken into account in future factor models. However, due to the homogeneous rating methods, a worldwide method for rating sustainability should be created, that fully represents the sustainability of the firms.

Bibliography

Armitage, S. (1995). Event study methods and evidence on their performance. Journal of economic surveys, 9(1), 25-52.

Awaysheh, A., Heron, R. A., Perry, T., & Wilson, J. I. (2020). On the relation between corporate social responsibility and financial performance. *Strategic Management Journal*, *41*(6), 965-987.

Bams D. & van der Kroft B. (2022). Tilting the Wrong Firms? How Inflated ESG Ratings Negate Socially Responsible Investing under Information Asymmetries.

Bauer, R., Koedijk, K., & Otten, R. (2005). International evidence on ethical mutual fund performance and investment style. *Journal of banking & finance*, *29*(7), 1751-1767.

Belghitar, Y., Clark, E., & Deshmukh, N. (2014). Does it pay to be ethical? Evidence from the FTSE4Good. *Journal of Banking & Finance*, *47*, 54-62.

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

Bollen, N. P. (2007). Mutual fund attributes and investor behavior. *Journal of Financial and Quantitative Analysis*, *42*(3), 683-708.

Bourke, M. (1997). Straining Out Gnats. *Christian Ethical Investment Group Newsletter*, *5*, 213-227.

Cao, J., Titman, S., Zhan, X., & Zhang, W. (2020). *ESG preference, institutional trading, and stock return patterns* (No. w28156). National Bureau of Economic Research.

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

Cortez, M. C., Silva, F., & Areal, N. (2012). Socially responsible investing in the global market: The performance of US and European funds. *International Journal of Finance & Economics*, *17*(3), 254-271.

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

Durán-Santomil, P., Otero-González, L., Correia-Domingues, R. H., & Reboredo, J. C. (2019). Does sustainability score impact mutual fund performance?. *Sustainability*, *11*(10), 2972.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, *33*(1), 3-56.

Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of political economy*, *81*(3), 607-636.

Geczy, C., Stambaugh, R. F., & Levin, D. (2005). Investing in socially responsible mutual funds. *Available at SSRN 416380*.

Gregory, A., Matatko, J., & Luther, R. (1997). Ethical unit trust financial performance: small company effects and fund size effects. *Journal of Business Finance & Accounting*, *24*(5), 705-725.

Hamilton, S., Jo, H., & Statman, M. (1993). Doing well while doing good? The investment performance of socially responsible mutual funds. *Financial analysts journal*, *49*(6), 62-66.

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

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk, econometrica, vol. 47.

Kempf, A., & Osthoff, P. (2007). The effect of socially responsible investing on portfolio performance. *European financial management*, *13*(5), 908-922.

Kreander, N., Gray, R. H., Power, D. M., & Sinclair, C. D. (2005). Evaluating the performance of ethical and non-ethical funds: a matched pair analysis. *Journal of Business Finance & Accounting*, *32*(7-8), 1465-1493.

Krivin, D., Patton, R., Rose, E., & Tabak, D. (2003). Determination of the appropriate event window length in individual stock event studies. *Available at SSRN 466161*.

Latane, H. A., & Jones, C. P. (1977). Standardized unexpected earnings--a progress report. *The Journal of Finance*, *32*(5), 1457-1465.

Lioui, A., & Tarelli, A. (2022). Chasing the ESG factor. *Journal of Banking & Finance*, *139*, 106498.

MacKinlay, A. C. (1997). Event studies in economics and finance. *Journal of economic literature*, *35*(1), 13-39.

Mallin, C. A., Saadouni, B., & Briston, R. J. (1995). The financial performance of ethical investment funds. *Journal of Business Finance and Accounting*, *22*, 483-483.

McCann, L., Solomon, A., & Solomon, J. F. (2003). Explaining the growth in UK socially responsible investment. *Journal of General Management*, *28*(4), 15-36.

McWilliams, A., & Siegel, D. (1997). Event studies in management research: Theoretical and empirical issues. *Academy of management journal*, *40*(3), 626-657.

Miller, A.: 1992, 'Green Investment', in D. Owen (ed.), Green Reporting: Accountancy and the Challenge of the Nineties (Chapman and Hall, London), pp. 242–255.

Moorman, C., & Lehmann, D. R. (2004). Assessing Marketing Strategy. *Marketing Science Institute, Cambridge MA*.

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

Renneboog, L., Ter Horst, J., & Zhang, C. (2011). Is ethical money financially smart? Nonfinancial attributes and money flows of socially responsible investment funds. *Journal of Financial Intermediation*, *20*(4), 562-588.

Riedl, A., & Smeets, P. (2017). Why do investors hold socially responsible mutual funds?. *The Journal of Finance*, *72*(6), 2505-2550.

Rockness, J., & Williams, P. F. (1988). A descriptive study of social responsibility mutual funds. *Accounting, Organizations and society*, *13*(4), 397-411.

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

Sparkes, R., & Cowton, C. J. (2004). The maturing of socially responsible investment: A review of the developing link with corporate social responsibility. *Journal of business ethics*, *52*(1), 45-57.

Committed to Transparency—People, Planet, Coffee. (n.d.). Retrieved from https://www.starbucks.com/responsibility/reporting-hub/

Vasudeva, G. (2013) Weaving together the normative and regulative roles of government: How the Norwegian sovereign wealth fund's responsible conduct is shaping firms' cross-border investments. *Organization Science*, 24: 1662–1682.

Yang, R. (2022). What do we learn from ratings about corporate social responsibility? New evidence of uninformative ratings. Journal of Financial Intermediation, 52, 100994.

Appendix

Figure A1 Simplified Morningstar Sustainability Rating Calculation Steps

Source: Morningstar/Sustainalytics

Table A1 Pairwise correlations

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
(1) Annual Return	1.000													
(2) Investor return	0.999	1.000												
(3) Corp. Sust. % Rank	0.056	0.083	1.000											
(4) MSR	-0.085	-0.105	-0.917	1.000										
(5) Asset Alloc. Equity %	0.076	0.098	-0.107	0.086	1.000									
(6) Net Expense Ratio	-0.110	-0.150	-0.046	0.037	0.047	1.000								
(7) Europe	-0.245	-0.231	-0.081	0.084	0.000	0.120	1.000							
(8) North America	0.266	0.259	0.086	-0.091	0.014	-0.138	-0.971	1.000						
(9) Asia	-0.102	-0.122	-0.026	0.038	-0.056	0.074	-0.113	-0.128	1.000					
(10) Fund size in Billion $	0.047	0.053	0.081	-0.046	-0.151	-0.118	-0.109	0.115	-0.028	1.000				
(11) Turnover Ratio	0.010	0.020	0.159	-0.158	0.010	0.001	-0.181	0.159	0.090	-0.113	1.000			
(12) Large Growth	-0.089	-0.142	-0.226	0.244	0.019	-0.029	0.167	-0.182	0.065	0.069	-0.175	1.000		
(13) Mid Blend	0.167	0.159	0.028	-0.036	-0.008	0.074	-0.012	0.017	-0.019	-0.051	0.059	-0.207	1.000	
(14) Small Value	0.190	0.210	0.077	-0.083	0.008	0.016	-0.129	0.134	-0.021	-0.031	0.059	-0.129	-0.048	1.000

Table A2 CAARs of high and low ESG firms for all event windows after a negative earnings surprise

	{T-20, T+20}		{T-5, T+5}		{T-3, T+3}	
EVENT DATE	CAAR low ESG	CAAR high ESG	CAAR low ESG	CAAR high ESG	CAAR low ESG	CAAR high ESG
-20	-0,25%	0,44%	-	-	-	-
-19	-0,88%	0,57%	-	-	-	-
-18	-0,98%	0,74%	-	-	-	-
-17	-1,16%	1,19%	-	-	-	-
-16	-1,98%	1,28%	-	-	-	-
-15	-2,33%	1,04%	-	-	-	-
-14	-3,27%	0,72%	-	-	-	-
-13	-3,79%	0,47%	-	-	-	-
-12	-4,14%	0,48%	-	-	-	-
-11	-4,41%	0,00%	-	-	-	-
-10	-4,34%	-0,08%	-	-	-	-
-9	-4,97%	-0,11%	-	-	-	-
-8	-5,07%	0,01%	-	-	-	-
-7	-5,11%	-0,02%	-	-	-	-
-6	-4,95%	-0,22%	-	-	-	-
-5	-4,79%	-0,13%	0,16%	0,09%	-	-
-4	-4,84%	-0,47%	0,11%	-0,25%	-	-
-3	-4,98%	-0,61%	-0,03%	-0,39%	-0,14%	-0,14%
-2	-5,43%	-0,90%	-0,48%	-0,68%	-0,60%	-0,43%
-1	-5,58%	-1,45%	-0,63%	-1,23%	-0,75%	-0,98%
0	-5,55%	-3,79%	-0,60%	-3,57%	-0,71%	-3,32%
1	-8,13%	-6,46%	-3,18%	-6,24%	-3,29%	-6,00%
2	-8,12%	-6,09%	-3,17%	-5,87%	-3,29%	-5,63%
3	-8,58%	-6,04%	-3,63%	-5,82%	-3,74%	-5,58%
4	-9,59%	-6,17%	-4,64%	-5,95%	-	-
5	-9,88%	-6,19%	-4,93%	-5,97%	-	-
6	-10,12%	-6,18%	-	-	-	-
7	-11,06%	-6,99%	-	-	-	-
8	-11,60%	-7,25%	-	-	-	-
9	-11,87%	-7,42%	-	-	-	-
10	-13,13%	-7,72%	-	-	-	-
11	-13,11%	-8,22%	-	-	-	-
12	-14,11%	-8,54%	-	-	-	-
13	-14,64%	-8,95%	-	-	-	-
14	-15,58%	-8,89%	-	-	-	-
15	-15,58%	-9,16%	-	-	-	-
16	-15,97%	-8,92%	-	-	-	-
17	-16,76%	-9,37%	-	-	-	-
18	-17,14%	-9,58%	-	-	-	-
19	-17,69%	-9,89%	-	-	-	-
20	-17,63%	-9,91%	-	-	-	-

Table A2 CAARs of high and low ESG firms for all event windows after a positive earnings surprise

	{T-20, T+20}		{T-5, T+5}		{T-3, T+3}	
EVENT DATE	**CAAR low ESG**	**CAAR high ESG**	**CAAR low ESG**	**CAAR high ESG**	**CAAR low ESG**	**CAAR high ESG**
-20	0,42%	-0,04%	-	-	-	-
-19	0,49%	-0,02%	-	-	-	-
-18	0,62%	0,05%	-	-	-	-
-17	0,70%	0,22%	-	-	-	-
-16	0,55%	0,24%	-	-	-	-
-15	0,47%	0,19%	-	-	-	-
-14	0,05%	-0,04%	-	-	-	-
-13	-0,35%	-0,06%	-	-	-	-
-12	-0,73%	-0,17%	-	-	-	-
-11	-0,89%	-0,38%	-	-	-	-
-10	-1,22%	-0,52%	-	-	-	-
-9	-1,38%	-0,66%	-	-	-	-
-8	-1,33%	-0,82%	-	-	-	-
-7	-1,65%	-0,95%	-	-	-	-
-6	-1,56%	-1,21%	-	-	-	-
-5	-1,69%	-1,27%	-0,13%	-0,06%	-	-
-4	-2,02%	-1,29%	-0,46%	-0,09%	-	-
-3	-2,08%	-1,35%	-0,52%	-0,14%	-0,06%	-0,05%
-2	-2,05%	-1,49%	-0,49%	-0,28%	-0,03%	-0,20%
-1	-2,06%	-1,68%	-0,50%	-0,47%	-0,04%	-0,38%
0	-1,26%	-1,05%	0,30%	0,16%	0,76%	0,25%
1	0,02%	-0,97%	1,58%	0,24%	2,04%	0,33%
2	0,14%	-0,99%	1,70%	0,22%	2,16%	0,30%
3	0,10%	-1,12%	1,66%	0,09%	2,12%	0,17%
4	-0,30%	-1,25%	1,26%	-0,04%	-	-
5	-0,79%	-1,47%	0,77%	-0,26%	-	-
6	-0,97%	-1,45%	-	-	-	-
7	-1,40%	-1,59%	-	-	-	-
8	-1,73%	-1,82%	-	-	-	-
9	-1,84%	-1,84%	-	-	-	-
10	-2,40%	-1,91%	-	-	-	-
11	-2,45%	-2,15%	-	-	-	-
12	-2,73%	-2,21%	-	-	-	-
13	-3,03%	-2,33%	-	-	-	-
14	-3,49%	-2,35%	-	-	-	-
15	-3,52%	-2,43%	-	-	-	-
16	-3,75%	-2,74%	-	-	-	-
17	-4,21%	-3,01%	-	-	-	-
18	-4,44%	-3,17%	-	-	-	-
19	-4,52%	-3,17%	-	-	-	-
20	-4,43%	-3,26%	-	-	-	-

Table A3 CAARs of high and low ESG firms for all event windows after a positive earnings surprise

	{T-20, T+20}		{T-5, T+5}		{T-3, T+3}	
EVENT DATE	CAAR low ESG	CAAR high ESG	CAAR low ESG	CAAR high ESG	CAAR low ESG	CAAR high ESG
-20	-0,25%	0,44%	-	-	-	-
-19	-0,88%	0,57%	-	-	-	-
-18	-0,98%	0,74%	-	-	-	-
-17	-1,16%	1,19%	-	-	-	-
-16	-1,98%	1,28%	-	-	-	-
-15	-2,33%	1,04%	-	-	-	-
-14	-3,27%	0,72%	-	-	-	-
-13	-3,79%	0,47%	-	-	-	-
-12	-4,14%	0,48%	-	-	-	-
-11	-4,41%	0,00%	-	-	-	-
-10	-4,34%	-0,08%	-	-	-	-
-9	-4,97%	-0,11%	-	-	-	-
-8	-5,07%	0,01%	-	-	-	-
-7	-5,11%	-0,02%	-	-	-	-
-6	-4,95%	-0,22%	-	-	-	-
-5	-4,79%	-0,13%	0,16%	0,09%	-	-
-4	-4,84%	-0,47%	0,11%	-0,25%	-	-
-3	-4,98%	-0,61%	-0,03%	-0,39%	-0,14%	-0,14%
-2	-5,43%	-0,90%	-0,48%	-0,68%	-0,60%	-0,43%
-1	-5,58%	-1,45%	-0,63%	-1,23%	-0,75%	-0,98%
0	-5,55%	-3,79%	-0,60%	-3,57%	-0,71%	-3,32%
1	-8,13%	-6,46%	-3,18%	-6,24%	-3,29%	-6,00%
2	-8,12%	-6,09%	-3,17%	-5,87%	-3,29%	-5,63%
3	-8,58%	-6,04%	-3,63%	-5,82%	-3,74%	-5,58%
4	-9,59%	-6,17%	-4,64%	-5,95%	-	-
5	-9,88%	-6,19%	-4,93%	-5,97%	-	-
6	-10,12%	-6,18%	-	-	-	-
7	-11,06%	-6,99%	-	-	-	-
8	-11,60%	-7,25%	-	-	-	-
9	-11,87%	-7,42%	-	-	-	-
10	-13,13%	-7,72%	-	-	-	-
11	-13,11%	-8,22%	-	-	-	-
12	-14,11%	-8,54%	-	-	-	-
13	-14,64%	-8,95%	-	-	-	-
14	-15,58%	-8,89%	-	-	-	-
15	-15,58%	-9,16%	-	-	-	-
16	-15,97%	-8,92%	-	-	-	-
17	-16,76%	-9,37%	-	-	-	-
18	-17,14%	-9,58%	-	-	-	-
19	-17,69%	-9,89%	-	-	-	-
20	-17,63%	-9,91%	-	-	-	-