ERASMUS UNIVERSITY ROTTERDAM
ERASMUS SCHOOL OF ECONOMICS
MSc Economics & Business
Master Specialisation Financial Economics

Mutual Fund Outperformance

Does sustainability come at a cost?

Author: N.G.B. Simon
Student number: 331927
Thesis supervisor: Dr. J.J.G. Lemmen
Second reader: Dr. J. Kil
Finish date: January 2020

PREFACE AND ACKNOWLEDGEMENTS

Firstly, I want to thank all my friends and family that supported me through this process. Especially for their patience! This thesis was a thorn in my side for longer than I would have wanted it to be, but it finally came to an end and I can happily say 'I'M DONE!'. I would also like to thank my supervisor Dr. Lemmen for his patience. This process has made me learn a lot of differing aspects of research, mutual funds and writing and I hope I get to use all my acquired knowledge in my future endeavours.

ABSTRACT

Following the modern portfolio theory, sustainable investing should lead to a less optimal portfolio due to it limiting its investment universe. A lot of literature has evaluated the performance of sustainable investing, especially as it has seen an immense growth over the years. One interesting finding is that over the years the interest in sustainability has been increasing and so has its performance. This paper tries to investigate if sustainable investing comes at the cost of financial risk-adjusted performance. Calculating Jensen's alpha as the first performance measure using the CAPM, the Fama-French three-factor model, the Carhart four-factor model, the market timing model and the Treynor and Mazuy model; with the Sharpe ratio as a secondary performance measure. A total of 1,140 global equity mutual funds will be analysed in this paper. The funds are divided into 3 categories, high sustainable mutual funds, low sustainable mutual funds and conventional mutual funds. The results show that for the CAPM the conventional funds slightly outperform the high and low funds. The other measures show no underperformance for the high funds, with the market timing model and Treynor and Mazuy model showing that the high funds outperform their conventional and low counterparts.

Keywords: Sustainable Investing, Hypothesis Testing, Time-Series Models, Ratings and Rating Agencies, Economic Cycles
JEL Classification: C12, C22, G23, G24, Q56,

TABLE OF CONTENTS

CHAPTER 1 Introduction

In this day and age, sustainability has become a prevalent issue. From office gossip to nation's politics, everyone is starting to engage on the issue of sustainability (Financial Times Ltd. (2012). This has led to an increase in awareness about sustainability. The growing awareness and interest for sustainability leads to new opportunities for everyone involved, including businesses, individual and institutional investors. One of the ways that investors invest in these new opportunities in sustainability is through mutual funds that focus on sustainable investments. The increase in interest in sustainability can also be seen in the investment universe where the net flows into sustainable investment funds have grown in recent years (Last Word Media UK Limited. (2018) and Dolvin et al. (2018)). The expectations are that this is not a temporary trend. This can be seen in the impact that the increase in interest in sustainability has on the investment universe. The world of investment has changed from one aiming only at financial performance to also looking at the impact investments have, both positive and negative, on the environment and society (US SIF (2019).

Since the rise in popularity of sustainable investing, there have been many interpretations of what a sustainable investment is. To minimize the room for misinterpretation, this paper will try to define the interpretations used within. Terms that are frequently used when referring to sustainable investments are ESG and SRI. ESG, in the world of investments, stands for Environmental Social and Governance. This term is mainly used as a manner of assessing the impact an investment has in the framework of sustainability. The second term SRI can stand for two things, namely Socially Responsible Investment or Sustainable and Responsible Investment (EUROSIF A.I.S.B.L. (2008). The latter encompasses the first and adds a sustainable dimension to it, which will be the way we interpret SRI throughout this paper. Later on, in the paper the definitions of these terms will be elaborated on in more detail.

To assess the extent of sustainability in the mutual funds within this research, the Morningstar database will be used. Morningstar sustainability rating and ESG scores are some of the variables that Morningstar analyses and presents in their database (Morningstar Inc. (2016). Morningstar is a respected and reliable source, when it comes to data on mutual funds (Morningstar Inc. (2018). The Morningstar Sustainability Rating is used to construct the mutual funds types within this research. This led to a total of 1,213 funds, which are then also filtered to reach our final total of 1,140 funds.

Being one of the foremost issues facing the world (United Nations (n.d.), there has been quite a bit of research done about sustainability, which is partly summarized in Ambec and Lanoie (2008). This is no different in the case of mutual funds and investing. There has been substantial research done on the performance of sustainable investing throughout the years, with a lot of differing results, which a non-

exhaustive list can be seen in Ito et al. (2013). At first glance, following the Modern Portfolio Theory (MPT) Markowitz (1952), limiting the investment universe to only sustainable investments would lead to a less optimal portfolio. On the other hand, studies have shown that the throughout the years the amount of options available as sustainable investments are increasing and becoming more and more lucrative. This can lead to all (or most) of the sustainable investment options, within a portfolio, to be on the Efficient Frontier of the Market Portfolio Theory, which wouldn't necessarily lead to a less optimal portfolio. Which leads to the research question if, due to the increase in sustainable investment opportunities, portfolios that consist of sustainable investments have seen an increase in performance throughout the years.

One of the ways to evaluate how different types of investment portfolios do empirically is through mutual funds. Mutual funds are investment vehicles made up of a pool of assets from investors, which is operated by a professional investor which attempts to maximize his or her portfolio. To do so, they typically invest in the most lucrative options available to them. The same as individual investors have different risk appetites, a mutual fund will also try to maximize its portfolio to fit its investment mandate(s). Because of this characteristic, varying types of mutual funds exist, which makes mutual funds an ideal way to see how well sustainable investments can financially perform. To do this, this paper will evaluate mutual funds that invest in sustainable assets/firms and compare them to other mutual funds in attempt to see if investing in sustainable assets/firms comes at a cost. This paper will conduct a comparative analysis similar to Ibikunle and Steffen (2015) in which 3 categories of funds will be researched and compared on their financial performance throughout time. The 3 categories in this research are high funds, low funds and conventional funds. The high funds category consists of the highest 10% of mutual funds rated on their sustainability and the low funds category consists of the lowest 10%. Lastly conventional funds are funds that have little to no limitations on where to allocate its resources, which in line with Markowitz (1952) would lead to an optimal portfolio if properly managed. The specifics of each of these categories of mutual funds will be further explained later on in this paper. Furthermore, a large time period will be used, namely 1990 to 2018. In such a large time period market cycles can also have an influence on the performance of mutual funds. Hence this research will also expand to compare the performance of the different categories of funds in the different market cycles.

To measure the financial performance of the different categories of mutual funds this paper will make use of several different methods. To evaluate the different categories of mutual funds we will compare them through risk-adjusted returns, using some of the most successful performance measures according to Climent and Soriano (2011). Firstly, the CAPM model will be used to see if, in the simplest case, high funds show outperformance, via a positive Jensen's alpha estimate, of the market and if this outperformance is greater than that of the low – and/or conventional funds. The other methods that will be used to compare the funds similar to the aforementioned method will be the Fama and French 3-factor model, the Carhart 4-factor model, the Treynor and Mazuy dual beta model and the market timing model. One more method

will be used to evaluate the performance of the mutual funds is the Sharpe Ratio. These methods will all be used to attempt at answering the question above, to see if high funds show dissimilar performance compared to the other categories of funds. Here the MPT can hold, showing that sustainable mutual funds show underperformance due to limiting their investment universe. This paper will contribute to the literature by using a more recent time period in its analysis of the performance of sustainable funds; also adding market timing into the analysis to see what impact this can have on the performance of the funds. Due to the large time period the funds will have gone through different economic cycles, which will also be incorporated to analyse the impact this has on the performance. This will happen in 3 steps. First the performance will be calculated, where after the market timing aspect will be added into the analysis and lastly the economic cycles will be analysed to bring us to our concluding results.

In short, this paper will try to answer the following question.

Main question: Does sustainable investing come at the cost of financial performance?

Hypothesis 1: Highly sustainable mutual funds' return and risk (performance) do not differ from other types of mutual funds.

Hypothesis 2: Differing economic cycles do not affect the return and risk (performance) of highly sustainable mutual funds.

Hypothesis 3: The return and risk (performance) of highly sustainable mutual funds is not affected by market timing.

The results show that for the CAPM the conventional funds show significant outperformance of 0.08% towards the market benchmark, while the high and low funds do not. This implies that the conventional funds outperform the high funds. For the three-factor model and the four-factor model no significant outperformance is found in any of the categories. When considering the economic cycles, namely the bull and bear market, the high funds show significant outperformance of the market benchmark in each cycle. While the low and conventional funds only show outperformance of the market benchmark in the bear market; and to a lesser extent than the high funds. This implies that the high funds, when considering the economic cycles, outperform the low and conventional funds. Also, when considering the market timing ability of the mutual funds, the high funds show once more a significant outperformance of the market benchmark. This time the conventional funds do not show significant outperformance to the benchmark, while the low funds do, at the 10% level, but once again to a lesser extent than the high funds. Thus, implying that the high funds outperform the conventional funds as well as the low funds. Regarding the

Sharpe ratio, the high, low and conventional funds show ratios of 0.29, 0.31 and 0.18, respectively. Suggesting that the high funds have neither the highest nor the lowest risk-adjusted performance.

T his paper consists of 5 Chapters. Following the introduction in Chapter 1, Chapter 2 will consist of the theoretical models used for this research and literature about sustainability and mutual fund performance. This entails a short explanation of the models themselves followed by how these models will be implemented within this paper. Where after past research will be analysed and evaluated in order to create a better understanding of this research and its relevance. Chapter 3 will explain de methodology and data used in this research. This will start of by going through the steps taken to gather the data used for this research. This will be done by elaborating on the articles and research that led to the choice of data, where after the databases for the data will be highlighted on the uses and conflict they provided while attempting to extract the data necessary. Afterwards the data extracted will be thoroughly described regarding the various methods used to filter through it, therefore leaving us with the final information that will be used for the research. To conclude this chapter the use of the final data will be discussed regarding the various methods employed to manipulate the data in order to get the required results. Evidently the results follow in Chapter 4. Here the results from all the models, namely the Jensen's alpha for CAPM, 3-factor model, 4-factor model, market timing model and the dual beta model, will be calculated and interpreted. After all the individual results are interpreted there will be a broader interpretation of all the results followed by a short summary and conclusion of this chapter. To conclude this paper, Chapter 5 will summarize all the previous chapters and make a conclusion in which the aforementioned question gets an answer and ideas for further research will be stated.

CHAPTER 2 Literature Review

2.1 Theoretical models

This chapter will elaborate on the theoretical models that will be used within this paper. This papers' research entails a comparative analysis of the risk-adjusted returns of several types of mutual funds. To achieve this some of the most successful performance measures will be used, according to Climent and Soriano (2011). This performance measure is the Jensen's alpha and the Sharpe Ratio, the former will be used within different models. These models are the Fama and French 3-factor model and the Carhart 4-factor model. An in-depth explanation of each of these measures will follow.

2.1.1 Jensen's alpha

Jensen's alpha is the first performance measure this paper is going to use in the comparative analysis of mutual funds. Jensen's alpha is a performance measure which is used to calculate the excess return of a portfolio (or an individual investment). It was first used in 1968 by Michael Cole Jensen as a performance measure to evaluate mutual fund managers Jensen (1968). It is based on the capital asset pricing model (CAPM) that was introduced by multiple researchers that were building on Markowitz's Modern Portfolio Theory. The researchers that worked independently towards the introduction of the CAPM are Treynor (1961, 1962), Sharpe (1964), Lintner (1965a, 1965b) and Mossin (1966). Which they took from a micro analysis to a market analysis of price formation for financial assets according to Nobel Prize press release (1990). Thus, the CAPM can be used to price a mutual fund portfolio. From the CAPM onwards, Jensen provided mathematical evidence in his paper to arrive at his so-called Jensen's alpha. See Jensen (1968) for the full clarification on how Jensen used the CAPM to create the Jensen's alpha. Which resulted in the formula that can be seen below.

It will be used in the comparative analysis to evaluate the abnormal return of each of the categories of mutual funds.This method is a simple approach to measure each strategies' performance, by estimating the relationship of the market returns to the risk free rate. This is called the market risk premium. This is adjusted by the fund's risk appetite, which is the *beta* of the fund. A beta greater or lesser than 1 means that the fund is, respectively, less or more risk averse than the market. A beta of 1 means that the fund is equally risky as the market. Then this should be subtracted from the fund's return minus the risk free rate, which will give Jensen's alpha.

The calculations to get Jensen's alpha will be done through regression analyses, which is a proces to estimate the relationship among variables. Jensen's alpha etc will be one of the outputs that the regression analyses gives, with a significant alpha meaning that it is very likely that the mutual fund exibits risk

adjusted outperformance. A positive alpha shows that the fund or portfolio of funds has had higher returns than was expected in the same period. The opposite is true for a negative alpha, *ceteris paribus*. These alphas are one of the outputs that are to be used in the comparative analysis.

2.1.2 Three-factor alpha

The second model, used to calculate the three-factor alpha, in the comparative analysis is the Fama-French three-factor model. Fama-French three-factor model is a model building on the CAPM, which describes the excess returns of a portfolio, or in this case a portfolio of mutual funds, using only one factor. Fama-French three-factor model builds on the CAPM by adding more factors to the model. As the name of the model states, three factors will be used by adding SMB and HML as the two new factors to the model. SMB is an abbreviation for **S**mall **M**inus **B**ig, which refers to the difference between, similar weighted-average book-to-market equity ratio portfolios, returns of small-cap and big-cap stock portfolios. This factor accounts for the difference in returns of small stocks and big stocks (Fama and French (1993).
HML is the abbreviation for **H**igh **M**inus **L**ow, which refers to the difference between, similar weighted-average size, returns of high and low book-to-market equity ratio portfolios. HML accounts for the difference in returns of high and low book-to-market firms. Firms with a high book-to-market ratio are known as value stocks and those with low book-to-market ratios are known as growth stocks (Fama and French (1993).

In the comparative analysis the above formula will be used to assess the performance of each of the categories of mutual funds. Similar to the previous method this method will attempt to make a more distinct measurable obeservation of each type of fund's performance. This method is a bit more complex than the first one encompassing the return on a diversified portfolio of small firm stocks minus those of big firm stocks and the difference between the returns on diversified portfolios of high and low book-to-market stocks in attempt to create a greater explanatory model in explaining diversified portfolio returns (Fama and French 1992, 1993)[1]. This model is arguably the most succesful asset pricing model empirically and is widely used by practitioners Nobel prize lecture Eugene Fama (2013). Seeing that there are two factors added to the model, the amount of betas in the model has changed to β_1, β_2 and β_3. The first beta (β_1) is the same as the beta in the CAPM model, which is the fund's risk appetite towards the market risk premium. The second beta (β_2) is an indicator of the fund's portfolio expected returns that is explained due to the size-factor. A higher beta translates into the diversification of the fund portfolio towards a higher proportion in small stocks. This equals to a proportion of the fund's performance being accreddited to small-cap firms. The third and last beta (β_3) is an indicator of the fund's portfolio expected returns that is explained due to the book-to-market factor. A higher beta translates into the diversification of the fund portfolio towards a higher proportion in value stocks. This equals to a proportion of the fund's performance being accreddited

[1] In Fama E. and French K. their model proved a higher R^2 within their sample. CAPM showed an R^2 of 0.69 compared to the lowest R^2 of the Fama-French three-factor model of 0.83.

to value stocks. The Fama-French three-factor model will also be put through regression analyses. With the results to be interpreted as stated above and the resulting alpha will be a 3-factor alpha.

Above the basic CAPM is altered into the Fama-French three-factor model through adding two extra factors, namely SMB and HML. Fama and French added these factors to explain diversified portfolios performance to a greater extent, in which they succeeded for their sample.

2.1.3 Carhart four-factor alpha

The third performance measure that will be used for the comparative analysis is the Carhart four-factor model. Similar to Fama-French three-factor model, this model is also a performance measure building on the CAPM. It does this with the help of the Fama-French three-factor model and adding an extra factor to the model. The additional factor is the momentum factor, which is done by capturing Jegadeesh and Titman's (1993) momentum anomaly by Carhart (1997). Carhart finds that the performance of the Fama-French three-factor model estimates on mutual funds are more precise than his model, but in general are not economically different from the CAPM estimates. With adding his momentum factor, Carhart's estimates frequently differ from the CAPM estimates due to the significant loadings on the one-year momentum factor Carhart (1997). In his paper Carhart also states that he employs his model to try and explain returns, leaving the risk-based interpretations to the reader. In Carhart (1997) he gives the abbreviation of PR1YR to his momentum factor, but there are also other abbreviations used to specify this factor, namely WML, UMD and MOM. To get rid of any confusion, this paper will use the abbreviation of MOM to address the momentum factor. MOM refers to the difference between equally weighted portfolios of the best past performers[2] and the worst past performers Jegadeesh and Titman (1993). Adding this factor, the Fama-French three-factor model gives the Carhart four-factor model, which can be seen below.

The above formula will be used in the comparative analysis to assess the performance each of the categories of mutual funds. Not unlike the other methods, the Carhart four-factor model will also be used to attempt to make a more distinct measurable observation of each type of fund's performance. The added factor, compared to the Fama-French three-factor model, adds another view on explaining the mutual fund's returns. Due to the fact that the Carhart four-factor model frequently being economically different, the model will add an additional perspective on the interpretation of the results Carhart (1997). The added factor in the model also has its own beta, β_4. This beta (β_4) is an indicator of the fund's portfolio expected returns that is explained due to the momentum factor. A higher beta translates into the diversification of the fund portfolio towards a higher proportion in one-year return momentum stocks. This equals to a proportion of the fund's performnce being accreddited to one-year momentum stocks Carhart (1997). The

[2] Performers are ranked on the basis of their returns.

Carhart four-factor model will be put through similar regression analyses, as the before mentioned methods. These results will be lead to a 4-factor alpha and multiple betas that will be interpreted as stated in the above paragraphs.

After this third model this paper has a model that uses CAPM to calculate Jensen's Alpha, a model that added on to the basic CAPM model to create the Fama-French three-factor model and the third model that builds further on top of the Fama-French three-factor model to create the Carhart four-factor model through adding an additional factor, namely MOM. Carhart added this factor to explain returns, even though Fama-French three-factor model is more precise. However the Carhart four-factor model frequently gives different results, due to significant loadings on the one-year momentum factor, which adds a different perspective on the results which could be relevant in this comparative analysis. The next method differs from the above mentioned ones, in the fact that it does not use CAPM or an altered version of it. It is a different way to examine a fund's risk-adjusted performance.

2.1.4 Treynor and Mazuy model and market timing model

The final models that will be used in this analysis are the Treynor and Mazuy model and the market timing model. The purpose of these models is to measure the performance of the mutual funds similar as to the models above, only in this case a distinction will be made between bull and bear markets. For long time periods markets go through different cycles and it can be advantageous to take this into account when looking at the performance of mutual funds, as this can have an effect on the performance of a mutual fund. Treynor and Mazuy (1966) study if investment managers anticipate major turns in stock market. In which they show that there is no statistical evidence of investment managers outguessing the market. This led to the creation of the Treynor and Mazuy model, which adds an extra variable to CAPM. The equation used in this research is derived from Santi Paramita (2015). In this research the Treynor and Mazuy model is also added to the Fama French 3-factor model and the Carhart 4-factor model.

Similar to the Treynor and Mazuy model, the market timing model also measures performance making a distinction between bull and bear markets. The market timing model measures the performance for the bull and bear markets separately. Which displays the results separately for a downturn market and an upturn market.

Thus, multiple models are used in this research to analyse Jensen's alpha. From the CAPM (1) that was used by Jensen to create Jensen's alpha (2); rewritten with the alpha as dependent variable follows in the formula below (3). Fama and French then added their SMB and HML variables to create their three-factor model (4); rewritten with the alpha as dependent variable follows in the formula below (5). Afterwards Carhart added the MOM variable in his four-factor model (6); rewritten with the alpha as dependent

variable follows in the formula below (7). Adding Treynor and Mazuy's dual beta factor to the model gives Treynor and Mazuy's four-factor model (8); rewritten with the alpha as dependent variable follows in the formula below (9).

$$E(R_j) = R_F + \beta j[\, E(R_j) - R_f \,] \quad (1)$$

$$R_{jt} - R_{Ft} = \alpha_j + \beta_j(R_{Mt} - R_{Ft}) + u_{jt} \quad (2)$$

$$\alpha_j = (R_{jt} - R_{Ft}) - \beta_j(R_{Mt} - R_{Ft}) - u_{jt} \quad (3)$$

$$R_{jt} - R_{Ft} = \alpha_{jt} + \beta_1(R_{Mt} - R_{Ft}) + \beta_2 SMB_t + \beta_3 HML_t + u_{jt} \quad (4)$$

$$\alpha_j = (R_{jt} - R_{Ft}) - \beta_j(R_{Mt} - R_{Ft}) - \beta_2 SMB_t - \beta_3 HML_t - u_{jt} \quad (5)$$

$$R_{jt} - R_{Ft} = \alpha_{jt} + \beta_1(R_{Mt} - R_{Ft}) + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 MOM_t + u_{jt} \quad (6)$$

$$\alpha_j = (R_{jt} - R_{Ft}) - \beta_j(R_{Mt} - R_{Ft}) - \beta_2 SMB_t - \beta_3 HML_t - \beta_4 MOM_t - u_{jt} \quad (7)$$

$$R_{jt} - R_{Ft} = \alpha_{jt} + \beta_1(R_{Mt} - R_{Ft}) + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 MOM_t + \beta_5(R_{Mt} - R_{Ft})^2 + u_{jt} \quad (8)$$

$$\alpha_j = (R_{jt} - R_{Ft}) - \beta_j(R_{Mt} - R_{Ft}) - \beta_2 SMB_t - \beta_3 HML_t - \beta_4 MOM_t - \beta_5(R_{Mt} - R_{Ft})^2 - u_{jt} \quad (9)$$

The following holds for the market timing model.

For Bear markets if $R_{Mt} \leq R_F$

For Bull markets if $R_{Mt} > R_F$

Where:

$E(R_j)$ denotes the expected return on a portfolio j

R_F denotes the risk-free rate

β_j denotes the beta

$E(R_M)$ denotes the expected return on the market portfolio

α_j denotes the intercept, managerial stock picking ability; also known as Jensen's alpha

u_j denotes the error term, idiosyncratic risk

addition of subscript t to denote the time interval

SMB denotes the SMB factor
HML denotes the HML factor
MOM denotes the MOM factor

2.1.5 Sharpe ratio

The second performance measure that will be used for the comparative analysis is the Sharpe Ratio. This performance measure isn't like the first three methods which rely on the CAPM as the base of their measures. Sharpe Ratio is a performance measure that evaluates the fund's risk-adjusted performance. This ratio was introduced by William F. Sharpe as an attempt establish a measurement and prediction of mutual fund performance. He did this using the recent works of that time, as there was considerable progress made in three areas, namely theory of portfolio selection, the theory of pricing of capital assets under the conditions of risk and the general behaviour of stock-market prices Sharpe (1966). In the paper he elaborated on the steps he took to come to his ratio. For the full explanation on how W. Sharpe arrived at his now well-known Sharpe ratio see Sharpe (1966). The formula for the Sharpe ratio can be seen below.

$$S_p = \frac{E[\,R_p - R_f\,]}{\sigma_p} \tag{10}$$

Where:

S_p denotes the risk-adjusted performance; also knows as the Sharpe ratio
R_p denotes the asset/fund return
R_f denotes the risk-free return
σ_p denotes the standard deviation
$E[R_p - R_f]$ denotes the expected excess return of the asset/fund

The paper states that the best portfolio will be the one giving the best boundary. Solving the above equation gives the boundary, and the best boundary is the one which gives the greatest ratio. The Sharpe Ratio will be used in the comparative analysis to evaluate the risk-adjusted return of each of the categories of funds. The Sharpe Ratio will also be added to our attempts to get a more distinct measurable observation of each type of fund's performance. The Sharpe Ratio is an uncomplicated method to measure the performance of the funds, by only needing a few simple variables. The variables needed to calculate the Sharpe Ratio are the returns of the fund, the risk-free rate and the standard deviation of the fund.

With these variables' calculations are made to get to the Sharpe Ratio. First the risk-free return is subtracted from the return of the fund, for each time period. Afterwards the expected excess return is calculated, which

gives us the result for the numerator. The denominator is the standard deviation (σ) of the return of the fund. Dividing the numerator and denominator gives the Sharpe Ratio. These calculations, to get the Sharpe Ratio, will be subjected to significance tests to evaluate the attained results.

Table 1: Meta table on theoretical models

Author(s)	*Method*	*Application*
Jensen (1968)	Jensen's Alpha	Performance measure based on the market
Treynor (1961, 1962)	CAPM	Estimation of expected returns based on the market
Sharpe (1964)	CAPM	Estimation of expected returns based on the market
Lintner (1965a, 1965b)	CAPM	Estimation of expected returns based on the market
Mossin (1966)	CAPM	Estimation of expected returns based on the market
Fama and French (1992, 1993)	Fama-French three-factor model	Estimation of expected returns based on the market, SMB and HML
Jegadeesh and Titman (1993)	Momentum anomaly	MOM
Carhart (1987)	Carhart's four-factor model	Estimation of expected returns based on the market, SMB, HML and MOM
Sharpe (1966)	Sharpe Ratio	Performance measure based on volatility
Treynor and Mazuy (1966)	Treynor and Mazuy dual beta model	Estimation of expected returns based on the market and economic cycles
Santi Paramita (2015)	Treynor and Mazuy dual beta model	Estimation of expected returns based on the market and economic cycles

2.2 Empirical papers

There have been a lot of studies on the mutual fund's performance, sustainable investing and a combination of both. This generates a substantial amount of literature to look through. Due to this, it is inevitable that within all of these studies differing interpretations and results arise. This chapter will sift through, some of, the literature and try to organize each of the thoughts, perspectives and results of the different authors. This will lead to a greater understanding of the extent of this research and the interpretations made herein.

2.2.1 Defining SRI and ESG

As stated in the introduction, when referring to sustainable investments the terms SRI and ESG are often used. The past decades have seen a rapid growth in sustainable investments around the world (Renneboog et al., 2008a). Due to this increase the use of the terms SRI and ESG have also been used more often. These terms are used to explain if an investment should be seen as a sustainable investment. Since these terms also try to define the extent of sustainability, interpretation will play a big role in how these terms will be applied. To reduce the extent of interpretation of these terms the next couple of paragraphs will try to clarify how these terms will be used in this paper.

A Forbes article from 11th of July 2018 named "The Remarkable Rise Of ESG" talks about the pace that ESG investing has grown since the word 'ESG' was coined in 2005 by Ivo Knoepfel in his report named "Who Cares Wins". ESG investing built on Socially Responsible Investment, which was the movement beforehand. One of the reasons for the growth is the increase in data and tools to be able to make ESG a more tangible stipulation. This has seen a steady improvement due to launch of the Global Reporting Initiative (GRI) in 2000, with 80% of the world's largest corporation using GRI standards. The article shows that in 2018 ESG investing is at an estimated $20 trillion in assets under management. Nowadays, reports on the growth of sustainable investing are not in short supply[3]. In Climent and Soriano (2011) the introduction to their paper on "Green and Good? The Investment Performance of US Environmental Mutual Funds" starts of saying that there is an increase in interest in socially responsible investments and point out that USSIF reports that one ninth of all assets under management in the US in 2007 was involved in socially responsible investing. Renneboog et al. (2008a) also speak on the global explosive growth of socially responsible investments, over the past decade looking from 2008. Almost a decade later, Dolvin et al. (2018) start their paper by pointing out that the expansion of socially responsible investing funds started slowly and as of recent years the pace quickened. They also state an example that the USSIF report (2016) reports that assets of funds that focus on sustainable investing have seen a 33 percent growth since 2014. Almost a decade after Renneboog et al. (2008a) mentioned the explosive growth of socially responsible

[3] A large literature presents the growth in sustainable investments, however only a small number of papers will be discussed in this section since they carry the same point across. A few more examples of literature that speak about this growth are Kreander et al. (2005), Ito et al. (2013), Muñoz et al. (2014) and Ibikunle and Steffen (2015).

investing, over the decade before 2008, the relevance of the growth is still a topic of discussion. The evolution of the sustainable investing is one not to take lightly. From socially responsible investing to ESG investing to SRI investing, all being terms used for sustainable investing. That shows that the importance of sustainable investing is increasing over the years along with the standards defining it. The following paragraphs will elaborate on the definitions for SRI and ESG.

Firstly, if an investment is SRI, as the name gives away, the investment is considered sustainable (and responsible). Theoretically speaking it is as simple as that; yet applying this in practice is an entirely different story. In practice the definition of SRI still differs in interpretations. The differing definitions, in Europe, stem from diverse cultural and historical interpretations in the different European member states. Nonetheless Eurosif's European SRI Study 2016 attempts to reach a definition for SRI. The definition for SRI that they use is "*Sustainable and Responsible Investment ("SRI") is a long-term oriented investment approach, which integrates ESG factors in the research, analysis and selection process of securities within an investment portfolio. It combines fundamental analysis and engagement with an evaluation of ESG factors in order to better capture long term returns for investors, and to benefit society by influencing the behaviour of companies*" (SRI-study-2016-HR SRI Paragraph after intro). Even after defining SRI, Eurosif did not impose the definition in their report, although it is in line with their original goal. This leads to hesitation in accepting that definition for the use within this paper.

Another yet simpler definition can be found in Renneboog et al. (2008a), which is on socially responsible investments. They state that socially responsible investments "*is an investment process that integrates social, environmental and ethical considerations into investment decision making*". This definition is not necessarily wrong; however, it does not incorporate all of the ESG factors. Governance is one of the ESG factors that is not mentioned in the above definition, and as stated before ESG factors are a commonly used measurement for sustainable investments. Also, another fact is that the paper was published over a decade ago, which also makes it a less reliable source for the latest definition than a more recent paper. For these reasons another definition for SRI is sought after. The definition that encompasses the way this paper looks at SRI and sustainable investments is found in Ibikunle and Steffen (2015). They state that sustainable investing or SRI as it is more popularly known is "*generally defined as an investment approach that considers environmental, social and responsible corporate governance criteria in order to yield long-term competitive financial utility, as well as favourable societal effect*". As can be seen from the definitions from each paper, they all incorporate the ESG factors to a certain extent in their definitions. The ESG factors lie in the essence of sustainability, turning a conventional investment opportunity into a sustainable investment one. Even so these factors are a measure of sustainability and determine the extent of sustainability, rather than if an investment is sustainable or not. This will be further clarified in the next paragraph.

Environmental, Social and Governance (ESG) factors are a measure for evaluating investments to the extent of their sustainability. Investments must incorporate all three ESG factors to be considered sustainable. These ESG factors are defined by varying ESG issues. There is no exhaustive list of ESG issues, and they are regularly linked to one and another. Nonetheless, the Charter Financial Analyst (CFA) Institute published "Environmental, Social, and Governance Issues in Investing"[4], in 2015, which presents a list of examples on what ESG issues are comprised of.

Environmental	Social	Governance
▪ Climate change and carbon emissions ▪ Air and water pollution ▪ Biodiversity ▪ Deforestation ▪ Energy efficiency ▪ Waste management ▪ Water scarcity	▪ Customer satisfaction ▪ Data protection and privacy ▪ Gender and diversity ▪ Employee engagement ▪ Community relations ▪ Human rights ▪ Labor standards	▪ Board composition ▪ Audit committee structure ▪ Bribery and corruption ▪ Executive compensation ▪ Lobbying ▪ Political contributions ▪ Whistleblower schemes

These issues (and more) are analysed to assess the extent of the sustainability of an investment. The importance of ESG can also be seen in the report "Who Cares Wins" provided by the United Nations. In the report they emphasize the importance of ESG and how it can benefit performance. They state that it is in the interest of all market actors to include ESG factors in their investment decisions, due to the fact that it will ultimately contribute to more stable and predictable markets. Further along they explain the benefits successful firms enjoyed ascribable to them managing the entire range of ESG issues relevant to their business, '*thereby achieving the best results in terms of value creation*'. The report made the benefits and results of incorporating ESG factors in investment decisions clearer. Also reporting on ESG factors improved. All of this made the interest on ESG performance rise, and with this rise came the increasing demand for data on these factors.

There are also a lot of data providers that provide sustainability and ESG measures of firms. This data makes it easier to determine to what degree a fund is sustainable. Different data providers have different approaches on the way they interpret and supply their data[5]. One of the more prominent data providers is Morningstar. They have been a popular data provider on asset management for many years and are

[4] Herein empirical analyses are done to disclose the importance of ESG, also in-depth clarifications can be found on the benefits and issues ESG currently faces. Even though these subjects are prevalent, they are not within the scope of our research and shall not be discussed.
[5] A few off these data providers can be found on the Harvard Law School Forum.
Link: https://corpgov.law.harvard.edu/2017/07/27/esg-reports-and-ratings-what-they-are-why-they-matter/

considered to be a powerful contender in the industry. Their ratings are often used by fund managers and they provide multiple ratings and scores on several sustainable and ESG matters, making it an ideal database for this paper. They provide a portfolio ESG score, which is an asset-weighted average of normalized company-level ESG scores provided by Sustainalytics. They also have a fairly new rating, which they introduced in 2016, called the Morningstar Sustainability Rating. The rating a fund gets is its normally distributed ordinal score and descriptive rank relative to the fund's global category [6]. These measures will be used to determine in which category a fund will be placed for the research of this paper.

SRI and ESG are terms that have seen immense growth over the past couple of decades along with the interest in sustainable investments. Through the various literature, discussed above, this paper tries to clarify the definitions and interpretations of SRI and ESG. Also described is how this paper intends to interpret these terms and use them within this research. The literature isn't restricted to defining SRI and ESG, and how they can benefit firms. The increase in interest also incited a lot of empirical research on the performance of mutual funds that incorporate SRI and ESG factors in their investment decisions. The next section will elaborate further on these researches.

2.2.2 Mutual fund performance analyses

The performance of mutual funds has been analysed and discussed throughout most of the 20th century, and in the 21st century this has not seen a reduction in interest. Due to the extent of interest in explaining mutual fund performance various methods for analysing mutual fund performance have been created. Some of these methods have been elaborated on in the previous chapter. This segment will take one step further and look at literature that attempts to explain mutual fund performance. The broader understanding for performance measures, attained from the last chapter, will help to comprehend the literature that analyse mutual fund performance. Although there are numerous methods to evaluate mutual fund performance, this paper will limit the literature to those that employ at least one of the methods from the previous chapter.

One of the earlier papers, namely Jensen (1968), analyses mutual fund performance in the period of 1945 to 1964. He gathers data from 115 mutual funds from Wiesenberger's *Investment* Companies. In this paper he introduces his performance measure, the Jensen's Alpha, which he then uses to measure the performance of mutual fund managers. Based on net returns Jensen shows that only 39 out of 115 show a positive α. The negative average value for α shows that on average the funds earn less per year than they should have, given their level of systematic risk. The results from Jensen (1968) show that mutual funds in his sample underperform, in the period of 1945 to 1964.

[6] Check Morningstar Inc. (2018) to see, in more detail, how the ESG score and sustainability rating are defined.

The next paper comes from William F. Sharpe, in which he introduced his Sharpe Ratio. In Sharpe (1966), Sharpe analyses 34 mutual funds in a period of 10 years, from 1954 to 1963. The results show that all mutual funds display a ratio of smaller than 1, with the lowest and highest ratio displaying a value of, respectively, 0.43116 and 0.77842. Two alternative interpretations are presented for these results. Which are, that the differences between mutual funds are explained by, either transitory or due to excessive expenditures or that the differences can be, at least partially, attributed to differences in management skill. The returns from 1944 to 1953 and the Dow-Jones Industrials are computed to attempt to explain these alternatives. The 1944 to 1953 period gives similar results to that of the 1954 to 1963 period. Comparing to the Dow-Jones Industrials, the mutual funds hold a portfolio that is on average at least as good as the Dow-Jones portfolio, but the results, after costs, fall short. In line with the findings of Jensen (1968) the funds show relatively poor performance for the similar periods used in their papers. Fama and French (1992, 1993) continue the analysis of mutual fund performance, using a period of July 1963 to December 1990. The Fama-French three-factor model originates from these papers. Although they focus on the explanatory role of the varying factors they implement, an α is calculated for each regression. Looking at the Fama-French three-factor model outputs for α shows an average of 0.08%. Whilst it is a positive result, it doesn't show a great extent of outperformance. Carhart (1997) makes his analysis on the period of July 1963 to December 1993, which is similar to Fama and French (1992, 1993). His results show only very slight evidence consistent with skilled or informed mutual fund managers, which is not unlike the results found in Fama and French (1992, 1993).

In Guercio and Reuter (2014) the following period is chosen, from 1993 to 2004. A lot of literature on mutual fund performance has been published before this paper, which gives this paper a broader foundation to base their research on. This foundation states that it is well known that the average actively managed mutual fund underperforms. The results from the earlier mentioned papers haven't specifically looked at actively managed mutual funds, but the results do not contradict that statement. Guercio and Reuter (2014) uses the four-factor alpha for its analysis and adds a direct sold dummy for funds that are directly sold to retail investors, in contrast to funds that are sold through intermediaries or brokers. There results show that there is little evidence that the actively managed funds underperform index funds, for the direct-sold segment and that the underperformance is limited to the broker-sold segment. The emphasis is on evidence for underperformance of the funds, in regard to the market, leaving outperformance not discussed. This can imply that there is little to no outperformance to speak of, which aligns with the results found in previous years.

Mutual funds have been thoroughly analysed over the years. One thing that the results persistently show is that the extent of mutual funds outperforming the market is negligible. The literature presented above has shown that, from different samples in the period of 1944 to 2004, this holds true. Outperforming the market is impractical, therefore outperforming your competitors should be the aim. This is also, most likely, the

reason why a large amount of literature focusses on the comparison of different types of mutual funds and/or strategies that are implemented by mutual funds. In accordance to this, and the methods of related literature, this paper will also compare the performance between various types of mutual funds.

2.2.3 Sustainable fund performance

To complete this chapter a look will be taken into sustainable fund performance. A great amount of literature can be found on sustainable fund performance, due to growing interest and importance of sustainability. The amount sustainable investment opportunities have also been growing, establishing a foundation for a new type of mutual fund, namely sustainable mutual funds. Limiting their investment universe to comprise of sustainable investments, while also maximizing performance. As stated above it is more sensible for a mutual fund to outperform its competitors, than it is the market. This is also how the following literature tries to measure the performance of sustainability.

One of the earlier studies on socially responsible investing is from Hamilton et al. (1993). Testing for outperformance in portfolios, they compare socially responsible portfolios to conventional portfolios. They gather data, from Lipper Analytical Services, for the period of January 1981 through December 1990 on 32 socially responsible funds, as identified by their managers. They then split these funds into two groups, one for funds established in 1985 or earlier and the second for funds established in 1986 or later. Then they grab a tenfold of conventional funds for each group to run their comparison on. Using Jensen's alpha, they measure the excess returns of the funds, using NYSE as their benchmark; and find that there is no significant outperformance in excess returns for socially responsible funds. They also find that there is no significant difference between the performances of socially responsible funds and conventional funds. Thus, showing that for the period of January 1981 through December 1990 socially responsible investments do not come at a cost in risk-adjusted returns. Of course, socially responsible investments are not the same as sustainable investments, but this research does give one of the first insights of funds that limit their investment horizon in a similar manner to the sustainable funds this paper is going to research. Sustainable investing further limits the investment horizon, but expectations of this paper are that the sustainable funds will not differ much from the socially responsible portfolios in Hamilton et al. (1993).

In Mallin et al. (1995) they compare UK ethical funds to their non-ethical matched pairs and to the market. Ethical funds in their sample are defined as funds that might have a policy to not invest in certain industries or countries, which are deemed as unethical, and/or a policy to invest in environmentally friendly firms. The criteria used to match, the non-ethical funds to the ethical ones, are based on fund size and the date the fund was formed. They gather data, from Finstat, on all the UK ethical investment funds over the period of 1986-1993; giving 29 ethical funds which are matched with 29 non-ethical funds. Using FTASA (Financial Times All Share Actuaries Index) as their benchmark to calculate the performance using the Sharpe ratio,

Treynor ratio and Jensen's alpha, they analyse the risk-adjusted returns of the different types of mutual funds. In their initial analysis they compare the mean excess returns, in which the ethical funds underperform both the market as their matched counterparts. For the risk-adjusted performance measures, ethical funds and non-ethical funds both underperform the market. Their results also show that ethical funds outperform their non-ethical matched pairs. Kreander et al. (2005) extend on the study done by Mallin et al. (1995), by doing a similar matched pair analysis in a European setting. Kreander et al. (2005) gather data, for 30 ethical and 30 non-ethical funds, over the period January 1995 to December 2001. The criteria used to match the funds are age, size, country, and investment universe. With their approach they also used country and investment universe specific benchmarks. They measure the performance using the Sharpe ratio, Treynor ratio, Jensen's alpha and a size adjusted Jensen's alpha. The funds show an overall underperformance to their respective benchmarks when compared through Jensen's alpha. When all risk-adjusted performance measures are considered, the fund's performances are similar to their benchmarks. The comparison between ethical funds and their matched non-ethical funds, show no significance for a difference in performance.

Renneboog et al. (2008b) analyse socially responsible funds and their conventional counterparts. They gather their data of nearly all ethical/SRI mutual funds around the world. The data is attained from CRSP, Datastream and Worldscope databases for the period 1991-2003. The performance is measured through Jensen's alpha and Carhart's four-factor alpha. Their results show that socially responsible investments funds in many European, North American and Asia-Pacific countries strongly underperform their domestic benchmark portfolios but show no significant underperformance to their matched conventional counterparts in most countries. For the socially responsible investments funds in France, Ireland, Sweden and Japan Renneboog et al. (2008b) their results show consistency with their underperformance hypothesis, wherein socially responsible investments funds underperform conventional funds and that a higher screening intensity reduces the performance of socially responsible investments funds.

The next study is from Climent and Soriano (2011). They do a comparative analysis on green funds, SRI funds and conventional funds. Green funds in their research are a group within SRI funds, which seek environmentally responsible investments as well. Their data is gathered through CRSP US Mutual fund database over the period 1987 through 2009. Instead of matched pairs, Climent and Soriano (2011) evaluate their green funds using matched samples of the SRI funds and conventional funds. In regard to the SRI funds, the green funds are matched against an equally weighted portfolio of two SRI funds with fund age and end-of-period fund size as the matching criteria. In regard to the conventional funds, the green funds are matched against equally weighted portfolio of four funds with fund age, end-of-period fund size and investment objective as the matching criteria. This gives a sample of 49 funds; 7 green, 14 SRI and 28 conventional funds. The performance measures used in their analysis are the Treynor ratio, Sharpe ratio, Jensen's alpha and Carhart's four-factor alpha. Differing benchmarks are used for the calculation of

Jensen's alpha and Carhart's four-factor alpha. Their results show that green funds show lower returns than that of the conventional funds in the sample period 1987 through 2001 and the full sample period. In the more recent sample period 2001 through 2009, green funds show adjusted returns that do not significantly differ from the SRI funds or conventional funds.

Ito et al. (2013) do a comparative analysis of SRI funds, environmentally friendly funds and conventional funds, which is similar to that of Climent and Soriano (2011). Ito et al. (2013) analyse multiple regions in their sample, which contains data obtained from Bloomberg that covers the period 2000 through 2009 for EU, US and Japan. By dividing this sample period into two, a long period (2000-2009) and a short period (2006-2009) they attempt to examine the effect of the financial crisis on the funds' performances. The shortage function approach developed by Briec and Kerstens (2009) is used as their performance measure. Their results show that SRI funds outperform conventional funds. Environmentally friendly funds, although they did not perform as well as the SRI funds, performed in manners equal or superior to conventional funds.

Another paper that incorporates financial crises is Muñoz et al. (2014). They evaluate the performance of varying forms of socially responsible mutual funds, with the main focus on green funds. Gathering data from Morningstar for their study they obtain data on 18 US green funds and 89 European green funds over the sample period January 1994 to January 2013. These funds are split into two types, the ones that invest globally and the others that invests domestically. Similar to Climent and Soriano (2011) they build matched samples for the conventional portfolios, using the investment objective, inception date and total net assets under management as their matching criteria. For each green fund the matched samples consist of three conventional funds similar in size. Creating samples of crises periods and non-crises periods they analyse these periods along with the full sample using the alpha from the combination of Treynor and Mazuy and Carhart models. The results show that green funds do not perform worse than the other socially responsible mutual funds. The US global conventional funds perform better than their green US global funds counterparts. This does not hold for crisis market periods, where there is no underperformance seen for green US global funds.

Ibikunle and Steffen (2015) aim their study towards a comparative analysis of European green, black and conventional mutual funds over the period 1991-2014. They screen their data primarily using Thomson Reuters EIKON fund screener. Due to the limitations of the fund screener they also manually filter through the funds to clean for same-class and multi-country listings, also to get to their final sample of green funds they manually look at each individual fund to assess if they belong within the green category. Then they extract the monthly return index data from Datastream. The funds are analysed through a matched pair analysis on age and size. This leaves them with 175 green funds, 259 black funds and 976 conventional

funds. Firstly, a single-factor CAPM with varying benchmarks[7] is used to evaluate the performance. Secondly, a Carhart's four-factor alpha is used to evaluate the performance in multiple ways[8]. Lastly, they add dummy variables for each of their analyses for the fund type. The results show that over the full sample period, the green mutual funds underperform their conventional counterparts. To their black counterparts there is no significant difference, for the full sample period. The green funds' performance increases over time until no significant difference is seen between them and the conventional funds, and that they significantly outperform the black mutual funds over the period 2012-2014.

The last paper that will be discussed in this section is Dolvin et al. (2018). Similar to our research, this study uses the Morningstar database for its sustainability scores and evaluates mutual funds' performance on their Morningstar sustainability rating. They collect their data on all U.S. domestic mutual funds for the period 2012-2016. They use different methods to analyse their data. Firstly, they divide their data through their sustainability rating in 3 samples, a sample of the bottom 20%, the middle 60% and the top 20%. Using Carhart's four-factor alpha they do a comparative analysis between these 3 samples. Afterwards they take a comparative analysis between funds that have a social mandate and those who do not, with the funds without social mandate being in the top 50% of the sustainability rankings. They also take a look at the fund flows, initially looking at the median fund flows, and then doing a multivariate regression. The results from their analyses is that the risk-adjusted performance does not significantly differ between the different samples. Also, that funds that have a mandate show more stable cash flows.

The studies explained above show different perspectives on the comparative analysis that will be done within this paper. From the studies evaluating socially responsible funds, to ethical funds, to sustainable funds and green funds, a part of the evolution of the conventional funds to sustainable funds can be seen. With, Dolvin et al. (2018), the most recent study discussed making use of the Morningstar database that, as of 2016, has a Sustainability rating and ESG scores. This gives a more tangible way to analyse sustainable mutual funds. A lot of studies elaborate on the increase in interest in sustainability and that sustainable assets have seen immense growth the growth in the investing universe and afterwards compare sustainable funds to their conventional counterparts. This paper will also attempt to answer *if there is a cost in risk-adjusted performance for sustainable investments*. With the help of the Morningstar database, this paper will add more recent data to try to add clarification to the performance of the different types of mutual funds. Also, the performance in different economic cycles will be brought to light.

[7] The benchmarks used in the single-factor CAPM regressions are the Kenneth R. French data library-sourced global market factor, the Stoxx Europe 600, the S&P Global Alternative Energy and the FTSE global small cap index.

[8] The factors used for the Carhart four-factor alpha are the Kenneth R. French data library-sourced global factors for the full sample period, the sub sample period 1991-2002, the sub sample period 2003-2014 and the Kenneth R. French data library-sourced European factors for the full sample period.

Table 2: Meta table on empirical papers.

Author(s)	*Time Period*	*Region*	*Method*	*Results*
Climent and Soriano (2011)	1987-2009	US	Treynor ratio, Sharpe ratio, Jensen's alpha, Carhart four-factor model	The results for green mutual funds show underperformance to conventional and SRIfunds, except for the sample period 2001-2009 in which no significant difference is found.
Dolvin et al. (2018)	2012-2016	US	Carhart's four-factor model, Multivariate regression	The results show no significant difference between the different rated mutual funds.
Ibikunle and Steffen (2015)	2000-2009	EU	Jensen's Alpha, Carhart four-factor model, Dummy variables in both models	The results for green mutual funds show underperformance to conventional funds and no significant difference to black mutual funds. As well as no significant difference to conventional funds in the period 2012-2014.
Jensen (1966)	1945-1964	US	Jensen's alpha	The results for mutual funds show a negative α, showing underperformance to the market.
Sharpe (1968)	1945-1963	US	Sharpe ratio	The results for mutual funds show a Sharpe ratio < 1, showing underperformance to the market; after costs.
Fama and French (1992, 1993)	1963-1990	US	Fama-French three-factor model	The results for mutual funds show an average α of 0,08%, showing almost no outperformance to the market.
Carhart (1997)	1963-1993	US	Carhart four-factor model	The results for mutual funds show very slight evidence to skilled or informed managerial outperformance to the market.
Guercio and Reuter (2014)	1993-2004	US	Carhart four-factor model	The results for mutual funds show underperformance to the market.
Hamilton et al. (1993)	1981-1990	US	Jensen's alpha	The results for socially responsible funds show no significant outperformance for to the market and no significant difference to conventional funds.
Mallin et al. (1995)	1986-1993	UK	Sharpe ratio, Treynor ratio, Jensen's alpha	The results for ethical mutual funds show an underperformance to the market and an outperformance to the non-ethical mutual funds.
Kreander et al. (2005)	1995-2001	UK	Sharpe ratio, Treynor ratio, Jensen's alpha, size adjusted Jensen's alpha	The results for ethical mutual funds show an underperformance to the market for Jensen's alpha, similar performance to the market for the other performance measures and no significant difference to non-ethical mutual funds.

Renneboog et al. (2008b)	1991-2003	World	Jensen's alpha, Carhart's four-factor model	The results for socially responsible investment funds show that many EU, NA and Asia-Pacific countries underperform the market and no significant underperformance to conventional mutual funds in most countries. The socially responsible investment funds in France, Ireland, Sweden and Japan show underperformance to conventional funds.
Ito et al. (2013)	2000-2009	EU, US, Japan	Briec and Kersten (2009) shortage function	The results for SRI funds show outperformance to conventional funds and environmentally friendly funds.
Muñoz et al. (2014)	1994-2013	US, EU	Combination of Treynor and Mazuy and Carhart model	The results for green mutual funds show no underperformance to other social funds.

CHAPTER 3 Data and Methodology

In this chapter the data will be used for the research will be discussed. First a short description of the databases used for extracting the data will be given. Thereafter the data itself will be discussed regarding the data that is to be extracted and the filtering process to get to that data.

The data is gathered through Morningstar and Datastream. Morningstar has over 200 thousand mutual funds in its database, which has been reduced to a total of 1,216 mutual funds after applying the search criteria.

3.1 Data

The first database used to gather data is Morningstar. Morningstar is a global financial services firm from Chicago. They are a popular and influential data provider in the asset management industry. Blake and Morey (2000) show that the Morningstar Ratings are not a worse predictor than alternative predictors, such as Sharpe ratios, Jensen's alpha and four-factor model alphas among others. While the Morningstar ratings are comparable to other risk-adjusted performance measures, they present their results in a five-star system. This is a ranking system where the top funds receive a five-star rating and the lesser performance will result in a lower number of stars up until the bottom funds that receive a one-star rating. This makes it easier to understand and use for a typical investor. In Del Guercio and Tkac (2008) they analyse the effect that Morningstar has on mutual funds. They do this by looking at Morningstar ratings over time and how this influences mutual fund flows. Their result shows evidence that the Morningstar ratings have an effect on mutual fund flows, in that that investors significantly punish funds whose performance drops below the top one-third of funds to a three-star rating. On the other side they also find a disproportionate positive response to an upgrade to the highest five-star rating. In 2016 they came up with the Morningstar Sustainability Rating, which they define as *"The Morningstar Sustainability Rating is a measure of how well the holdings in a portfolio are performing on environmental, social and governance, or ESG, issues relative to a portfolio's peer group"*. Each fund that fits the Morningstar's requirements gets this rating. This rating is comprised of a three-step process. This process first calculates the Morningstar Portfolio Sustainability Score. Afterwards, the score the Morningstar Historical Portfolio Sustainability Score is calculated. Lastly, the Morningstar Sustainability Rating is assigned based on a portfolios' Morningstar Historical Portfolio Sustainability Score relative to its Morningstar Global Category.

From Morningstar the funds used in this research are extracted. First the entire mutual fund database is filtered on only including funds that have an inception date before the first of January of 2014, that invest globally, that invest in equity and they must have a Morningstar Sustainability Rating.

Through the Morningstar Sustainability Rating, which is explained above, the mutual funds will be categorized to their extent of sustainability. The Morningstar Sustainability Rating is comprised of 5 ratings of sustainability, namely 'High', 'Above Average', 'Average', 'Below Average' and 'Low'. The total of mutual funds that is attained through Morningstar is then filtered and transformed to be used in the next database. Similarly to Dolvin et al. (2017), this research assumes that each fund's sustainability score is constant prior to their first entry in the Morningstar Sustainability Rating.

Table 3: Morningstar data

Criteria	Restriction
Global Broad Category Group	= Equity
Investment Area	= Global
Inception Date	< 1-1-2014
ISIN	Not=
Morningstar Sustainability Rating	Not= NA
Fund Size	Not= NA
Return data	Available
Outliers	Increase to > 1000% or decrease to < 20%

The ISIN codes of the mutual funds gathered from Morningstar are used to gather data in Datastream. Datastream is a database provided by Thomson Reuters and it is one of the world's most comprehensive financial historical databases. They cover a wide range of data including the returns of mutual funds. The total of mutual funds from Morningstar is run through Datastream to extract return data for the time period July of 1990 to December 2018. For some mutual funds no data was found on Datastream, thus they are excluded from the total. Afterwards other measures of filtering were used for the total, such as excluding outliers. Outliers are funds that have seen an increase greater than 10 times their initial return and funds that have seen a decrease greater than 5 times their initial return.

3.2 Methodology

The methods that will be used to analyse the mutual funds' performance are the Jensen's alpha, Fama-French three-factor alpha, Carhart's four-factor alpha and the Sharpe ratio. Mutual fund portfolios will be split into three categories. The total for each portfolio consists of mutual funds that invest globally in equity and are active for more than 5 years. The total is divided in three categories using the Morningstar Sustainability Rating, which can be seen below.

Morningstar Sustainability Rating

Distribution	Score	Descriptive Rank	Rating Icon
Highest 10%	5	High	
Next 22.5%	4	Above Average	
Next 35%	3	Average	
Next 22.5%	2	Below Average	
Lowest 10%	1	Low	

Source: Morningstar, Inc.

The categories that are created are a High sustainability rating category, a Low sustainability rating category and a Conventional portfolio category. Based on the Morningstar sustainability rating these categories are, respectively, divided by the funds with a five-star rating, the funds with a one-star rating and the funds that have a two to four-star rating.

The theoretical models, that are going to be used in this research, have already been explained above. Thus, this chapter will jump to the implementation of these models within this research. Each of the models discussed above are used to evaluate mutual fund performance, which will be used to come to the risk-adjusted performance for each of our categories. This way the outperformance for each category can be quantified. Lastly, a paired t-test will be performed to see if there is a significant difference between the categories. The results shall be interpreted individually as well as together as a whole.

CHAPTER 4 Results

This chapter will elaborate on the results attained through the data and methodology. Firstly, descriptive statistics with each categories' Sharpe ratio will be shown, where after each model will have its results shown and they will be individually interpreted. Afterwards a combined interpretation of all the results will be presented.

First the mean return and standard deviation are calculated, which are then used to calculate the Sharpe Ratio. Where after Jensen's alpha will be calculated for all the theoretical models. After every table in interpretation will be given on the results.

Table 4: Descriptive Statistics and Sharpe Ratio

This table provides summary statistics on all three portfolio categories. The average portfolio returns are calculated for each month based on an equally weighted portfolio of all funds and then averaged out over the whole time period (July 1990 to December 2018). Both the *"Mean Return (%)"* as well as the corresponding *"Standard Deviation (%)"* are annualised. The average risk-free rate, which is used for the Sharpe Ratio is taken from the *Kenneth R. French data library*, for the period is 2.65%

Sustainability rating	*Mean Return (%)*	*Standard Deviation (%)*	*# Funds*	*Sharpe Ratio*
High	7.03	14.62	112	0.29
Low	7.94	16.21	68	0.31
Conventional	5.28	14.14	960	0.18

Table 4 shows the descriptive statistics and the Sharpe ratio for each of the three mutual fund categories. The funds a Low sustainability rating are the smallest group in the analysis with 68 funds in the category. The category with a high sustainability rating has 112 funds in the total, leaving the conventional category with 960 funds. From July 1990 to December 2018 the high sustainable mutual funds show an average annualised return of 7.03%, which is higher than the 5.28% for the conventional category but lower than the 7.94% for the low category. High funds that limit their investment horizon are still able to show a higher annualised return than their conventional counterparts.

The conventional funds have the lowest annualised standard deviation (14.14%), which can be explained by the less restrictive nature of their investments. Since they do not restrict themselves to invest in specific sustainable investments, they can enjoy the benefits of full portfolio diversification; which may entail both highly sustainable as low sustainable investment options. On the contrary the low funds are the riskiest with an annualised standard deviation of 16.21%, with the high funds in between these two with an annualised standard deviation of 14.62%.

Our first performance measure, the Sharpe ratio, is also shown in table 4. After seeing the annualised return and standard deviation it is not surprising to see that the high funds have a Sharpe ratio of 0.29, which is in between the low and conventional funds that have Sharpe ratios of 0.31 and 0.18, respectively. This shows that the low funds show the best performance within this analysis, with the high funds being a close second and the conventional funds at the bottom for the period July 1990 to December 2018.

Overall table 4 shows that there isn't a big difference between the high funds and low funds. Mainly the low funds are riskier, which can be seen by the higher standard deviation, but they do have a higher return for this higher risk resulting in a higher Sharpe ratio than the other funds. The high funds come a close second to the low funds regarding all results. And both show higher returns and Sharpe ratios than the conventional funds.

Table 5: Empirical results for one-factor (CAPM) Regressions using the *Kenneth R. French data library-sourced* global market factor

This table reports the regression results, with robust standard errors, of the monthly CAPM using a global market factor. The global MKT factor portfolio from the *Kenneth R. French data library* is used as a market proxy to measure the risk-adjusted returns of each of the portfolio categories. α measures of the risk-adjusted abnormal return relative to the market proxy. β_{MKT} measures the effect of the MKT factor. The t-statistics are depicted in parentheses. *, ** and *** correspond to statistical significance at 10%, 5% and 1% levels respectively. The time period is July 1990 to December 2018.

Sustainability rating	α	β_{MKT}	*Adj.* R^2
High	0.02 (0.15)	0.75*** (21.17)	0.57
Low	-0.02 (-0.12)	0.91*** (28.45)	0.71
Conventional	0.08*** (2.83)	0.83*** (32.38)	0.76

Table 5 presents the CAPM results for each of the three mutual fund categories using the *Kenneth R. French data library-sourced* global factor as the market benchmark. The global market benchmark is used since the mutual funds within this analysis are funds that invest globally. The first observation from the results show that the high and low funds show no significant outperformance of the market benchmark. On the other hand, the conventional funds do show a significant outperformance of the market benchmark. The 0.08% monthly outperformance of the market for the conventional mutual funds is significant at the 1% level. Hence, the conventional funds show the highest alpha and is also the only significant one which implies that the conventional funds have the better performance regarding these results. All the funds show a significant beta at the 1% level, with betas of 0.75, 0.91 and 0.83 for High, Low and Conventional, respectively. The high funds show the lowest beta meaning that they are the least sensitive to market risk in this analysis. Having the highest beta, the low funds are most sensitive to market risk. This implies a high correlation with the market and, as can be seen from the standard deviation in Table 4, heightened risks due to greater volatility. The funds that fit the model the best are the conventional funds, with an *adjusted* R^2 of 0.76. This is expected since the conventional funds are the most diversified, because their

portfolios do not consist of either heavily high sustainable investments on the one hand or low sustainable investments on the other hand.

Table 6: Empirical results for the Fama-French three-factor model Regressions using the *Kenneth R. French data library-sourced* global market factor, SMB and HML

This table reports the regression results, with robust standard errors, for the Fama-French three-factor model using a global market factor. The global factor portfolios from the *Kenneth R. French data library* are used as factors to measure the risk-adjusted returns of each of the portfolio categories. α measures of the risk-adjusted abnormal return relative to the market proxy. β_{MKT}, β_{SMB} and β_{HML} measure the effects of the MKT, SMB and HML factors. SMB corresponds to the return spread between a small cap portfolio and a large cap portfolio and HML is the difference in return between a value stock portfolio and a growth stock. The t-statistics are depicted in parentheses. *, ** and *** correspond to statistical significance at 10%, 5% and 1% levels respectively. The time period is July 1990 to December 2018.

Sustainability rating	α	β_{MKT}	β_{SMB}	β_{HML}	*Adj.* R^2
High	0.06 (0.40)	0.74*** (20.78)	0.21*** (2.82)	-0.13** (-2.00)	0.58
Low	0.12 (1.00)	0.89*** (28.68)	0.28** (2.13)	-0.32*** (-3.37)	0.74
Conventional	-0.11 (-1.02)	0.82*** (31.98)	0.12** (2.11)	-0.15*** (-3.19)	0.77

Table 6 presents the Fama-French three-factor model results for each of the mutual fund categories using the *Kenneth R. French data library-sourced* global factor as the market benchmark. Firstly, the alphas show no significant results. Which means that none of the fund categories show a performance that is significantly different than the market. There is no drastic change to be seen as to the betas of the market, the values remain similar and are significant at the 1% level for all three fund categories. For the factor estimations, both high (β_{SMB} = 0.21) as well as low funds (β_{SMB} = 0.28) show a decent exposure to small cap stocks; while the conventional funds (β_{SMB} = 0.12) show a small exposure to the small cap stocks. For the high funds this is significant at the 1% level, while for the low and conventional funds the result show significance at the 5% level. As discussed before, the sustainability movement has seen an increase as of late. This leads to more sustainable investing options arising in the market, skewing the amount of sustainable stocks more towards the small cap, which is in line with the results.

The high funds (β_{HML} = -0.13) show a lower amount of exposure towards growth stocks than the low (β_{HML} = -0.32) and conventional funds (β_{HML} = -0.15), with the low funds leaning the most towards them. For the high funds this is significant at the 5% level, while for the low and conventional funds the results are significant at the 1% level. These β_{HML} results show that in this analysis all three categories have a relatively small tendency towards growth stocks. Lastly, for the *adjusted* R^2 a small increase can be seen for all three categories, even though this increase isn't much it does show that the Fama-French three-factor model adds explanatory value to the analysis.

Table 7: Empirical results for the Carhart four-factor model Regressions using the *Kenneth R. French data library-sourced* global market factor, SMB, HML and MOM.

This table reports the regression results, with robust standard errors, for the Carhart four-factor model using a global market factor. The global factor portfolios from the *Kenneth R. French data library* are used as factors to measure the risk-adjusted returns of each of the portfolio categories. α measures of the risk-adjusted abnormal return relative to the market proxy. β_{MKT}, β_{SMB}, β_{HML} and β_{MOM} measure the effects of the MKT, SMB, HML and MOM factors. SMB corresponds to the return spread between a small cap portfolio and a large cap portfolio, HML is the difference in return between a value stock portfolio and a growth stock and MOM is the difference between a portfolio of the past 12 months' winners and a portfolio of the past 12 months' losers. The t-statistics are depicted in parentheses. *, ** and *** correspond to statistical significance at 10%, 5% and 1% levels respectively. The time period is July 1990 to December 2018.

Sustainability rating	α	β_{MKT}	β_{SMB}	β_{HML}	β_{MOM}	*Adj.* R^2
High	0.10 (0.63)	0.76*** (20.70)	0.13* (1.84)	-0.15** (-2.36)	-0.00 (-0.02)	0.59
Low	0.07 (0.50)	0.90*** (27.49)	0.20*** (2.92)	-0.31*** (-5.44)	0.05 (1.39)	0.74
Conventional	-0.08 (-0.70)	0.83*** (27.76)	0.11 (1.49)	-0.16*** (-2.67)	0.01 (0.22)	0.77

Table 7 presents the Carhart four-factor model results for each of the mutual fund categories using the *Kenneth R. French data library-sourced* global factor as the market benchmark. Similarly, to the Fama French three-factor model the alphas show no significant results, meaning that none of the fund categories show a performance that is significantly different than the market. The betas of the market remain significant at the 1% level for all three fund categories with values of 0.76, 0.90 and 0.83 for high, low and conventional, respectively. Again, the high fund category shows the least sensitive to market risk; followed by the conventional funds and then the low funds. This is in line with our two previous results, the CAPM and Fama-French three-factor model. Table 7 shows differing results as compared to table 6 when it comes to the exposure to small cap stocks in the three fund categories. All fund categories show a lower result, meaning less exposure to small cap stocks. This is only significant at the 10% level for the high funds ($\beta_{SMB} = 0.13$), while being significant at the 1% level for the low funds ($\beta_{SMB} = 0.20$) and the conventional funds ($\beta_{SMB} = 0.11$) showing no significance for the small cap stock exposure. For the high funds it is only significant at the 10% level, but it does show contradicting results to table 6. Showing in this model that the high funds exposure to small cap stocks is rather small itself.

All funds still show significant exposure towards growth stocks, with the high funds ($\beta_{HML} = -0.15$) being significant at the 5% level and the low ($\beta_{HML} = -0.31$) and conventional funds ($\beta_{HML} = -0.16$) at the 1% level. For the momentum factor, that is added in this model, no significant results are shown for none of the fund categories. Which implies that there is no momentum factor in which the past winners will keep winning in the short term or the other way around. The *adjusted* R^2 showed a slight increase for only the high funds, going from 0.58 to 0.59, and for the low and conventional funds no increase is shown. This leads to the conclusion that the momentum factor did not necessarily add more explanatory value into the results, except for a slight increase for the high funds.

Table 8: Empirical results for the market timing model Regressions using the *Kenneth R. French data library-sourced* global market factor

This table reports the regression results, with robust standard errors, for the market timing model using a global market factor. The global factor portfolios from the *Kenneth R. French data library* are used as factors to measure the risk-adjusted returns of each of the portfolio categories. α measures of the risk-adjusted abnormal return relative to the market proxy in Bull and Bear markets. Bull- and bear markets are defined as a time period wherein the market shows, respectively, positive- or negative market returns. β_{MKT}, β_{SMB}, β_{HML} and β_{MOM} measure the effects of the MKT, SMB, HML and MOM factors. SMB corresponds to the return spread between a small cap portfolio and a large cap portfolio, HML is the difference in return between a value stock portfolio and a growth stock and MOM is the difference between a portfolio of the past 12 months' winners and a portfolio of the past 12 months' losers. The t-statistics are depicted in parentheses. *, ** and *** correspond to statistical significance at 10%, 5% and 1% levels respectively. The time period is July 1990 to December 2018.

Sustainability rating	α	β_{MKT}	β_{SMB}	β_{HML}	β_{MOM}	*Adj.* R^2
High (Bull)	0.65** (2.02)	0.58*** (6.42)	0.15 (1.47)	-0.10 (-1.06)	0.05 (1.20)	0.20
Low (Bull)	0.25 (0.83)	0.85*** (9.74)	0.26 (1.54)	-0.27** (-2.00)	0.15* (1.67)	0.41
Conv. (Bull)	0.26 (1.16)	0.72*** (11.02)	0.10 (1.17)	-0.12 (-1.43)	0.05 (1.22)	0.45
High (Bear)	0.73** (2.58)	0.88*** (13.32)	0.13 (1.17)	-0.22*** (-2.64)	-0.18*** (-2.87)	0.59
Low (Bear)	0.60** (2.40)	0.96*** (16.98)	0.34*** (3.13)	-0.36*** (-4.32)	-0.20*** (-3.49)	0.73
Conv. (Bear)	0.48** (2.05)	0.94*** (17.20)	0.13 (1.43)	-0.20*** (-2.96)	-0.14*** (-2.68)	0.72

Table 8 presents the market timing model results for each of the mutual fund categories using the *Kenneth R. French data library-sourced* global factor as the market benchmark. For the bull market the low and conventional categories show no significant outperformance to the market benchmark. For the high category a monthly outperformance of 0.65% is shown, significant at the 5% level. This implies that the high category of funds has a better performance than the low or conventional categories. All three categories show significant results, at the 1% level, for the market beta. The results show market betas of 0.58, 0.85, 0.72 for the high, low and conventional categories, respectively. This means that the high category is the least sensitive to the market, the low category is the most sensitive to the market and the conventional category is in-between. Noteworthy is the adjusted R^2's of all the categories of funds are all lower, when compared to the other models. Hence, the market timing model for the bull market does not explain much of the returns of the mutual funds in this analysis. The high category of funds shows superior stock picking ability, in bull markets, compared to the other categories, but the results are subject to low explanatory power.

Similar results can be seen for the bear market, wherein the high category shows the highest measure of outperformance. Although this time every category shows alphas significant at the 5% level, with alphas of 0.73, 0.60 and 0.48 for the high, low and conventional funds, respectively. As opposed to the bull market all three categories of funds show a higher sensitivity towards the market, at the 1% significance level. The adjusted R^2 are also higher than for the bull markets, which implies that these results explain more of the data in the analysis. Similar to the bull market, the bear market also shows that the high category of funds shows superior stock picking ability, but the bear market results have higher explanatory power.

Table 9: Empirical results for the Treynor and Mazuy model Regressions using the *Kenneth R. French data library-sourced* global market factor

This table reports the regression results, with robust standard errors, for the Treynor and Mazuy model using a global market factor. The global factor portfolios from the *Kenneth R. French data library* are used as factors to measure the risk-adjusted returns of each of the portfolio categories. α measures of the risk-adjusted abnormal return relative to the market. β_{MKT}, β_{SMB}, β_{HML} and β_{MOM} measure the effects of the MKT, SMB, HML and MOM factors. SMB corresponds to the return spread between a small cap portfolio and a large cap portfolio, HML is the difference in return between a value stock portfolio and a growth stock and MOM is the difference between a portfolio of the past 12 months' winners and a portfolio of the past 12 months' losers. The t-statistics are depicted in parentheses. *, ** and *** correspond to statistical significance at 10%, 5% and 1% levels respectively. The time period is July 1990 to December 2018.

Sustainability rating	α	β_{MKT}	β_{SMB}	β_{HML}	β_{MOM}	B_5	*Adj.* R^2
High	0.37** (2.15)	0.73*** (18.12)	0.12 (1.48)	-0.15** (-2.28)	-0.02 (-0.42)	-0.01*** (-2.85)	0.61
Low	0.26* (1.71)	0.89*** (22.55)	0.25* (1.95)	-0.31*** (-3.07)	0.04 (0.54)	-0.01* (-1.93)	0.74
Conventional	0.12 (0.95)	0.81*** (26.76)	0.09 (1.33)	-0.16*** (-2.64)	-0.00 (-0.10)	-0.01** (-2.50)	0.78

Table 9 presents the Treynor and Mazuy model results for each of the mutual fund categories using the Kenneth R. French data library-sourced global factor as the market benchmark. The conventional category shows no significant outperformance of the market, whereas the high and low categories do show significant outperformance at the 5% and 10%, respectively. The high category shows the greatest performance with a monthly outperformance of 0.37%; whereas the low category shows a monthly outperformance of 0.26%. For the market beta, all categories show significant results, at the 1% level. The high funds show the lowest sensitivity towards the market with a beta of 0.73; with the low category showing the highest sensitivity of 0.89 and the conventional category with 0.81. The dual betas show a negative result of -0.01 for all categories of funds, which implies that the managers of the funds do not have the ability to time the market. The high, low and conventional categories of funds are significant as well, respectively at the 1% level, the 10% level and the 5% level. No big changes are found with the addition of the dual beta. The high category of funds shows a higher level of outperformance compared to the other categories of funds and the dual beta shows that none of the categories have the ability to time the market.

The market timing model and the Treynor and Mazuy model both were also analysed with just Fama-French factors as well as just the CAPM. These models show similar results to the ones presented above and don't contradict the results found above. The results of the other models can be seen in the appendix.

Table 10: Differences between mutual funds

This table reports the paired t-test between the different portfolio categories. The t-statistics are depicted with *, ** and *** corresponding to the statistical significance at 10%, 5% and 1% levels respectively. The time period is July 1990 to December 2018.

Pairs	*T-value*
High = Low	-0.50
High = Conventional	1.72*
Low = Conventional	-2.17**

Table 10 shows a simple paired t-test between the three mutual fund categories. This simple test shows if there is a significant difference between the fund categories. What can be seen in Table 8 is that there is no significant difference between the high and low funds. On the other hand, both the high funds and the low funds differ significantly from the conventional funds on a 10% and 5% level, respectively.

With all of the results given above an attempt is made at answering the research question, "*Does sustainable investing come at the cost of financial performance?*". To answer this question a comparison is made between funds with high sustainability rating, those with a low sustainability rating and those in between high and low. Firstly, the descriptive statistics show that the high funds have a higher annualised return than the conventional funds, but lower than the low funds. Also, the Sharpe ratio shows that a higher ratio for the high funds than the conventional funds, with again them being lower than the low funds. For the other models, CAPM, Fama-French three-factor model and Carhart's four-factor model, no significant result was found for the alphas; except for the Jensen's alpha for the conventional funds. Nonetheless, there is no evidence to show that sustainable investing comes at the cost of financial performance. For the market timing model in the bull market, the high funds show a significant outperformance of the market benchmark. For the bear markets all funds show significant outperformance, with the high funds showing the higher result. For the Treynor and Mazuy model the high funds also show a significant result that is higher than its counterparts. These models show that high funds outperform their counterparts, when keeping in mind the timing of the market. Table 10 also showed a significant difference between high and conventional funds. With that in mind, looking at table 4 it can be said that the high funds show outperformance compared to their conventional counterparts with a higher annualised return and Sharpe ratio. As for the high funds compared to the low funds, no significant difference is found in table 10 and in table 4 they also show similar Sharpe ratios. Taking everything into consideration the results show that sustainable investing does not necessarily come at a cost of financial performance and in some cases might even outperform.

CHAPTER 5 Conclusion

This research is a comparative analysis of the performance of three mutual fund categories, namely the high mutual funds, the low mutual funds and the conventional mutual funds. The high funds category consists of the highest 10% of mutual funds rated on their sustainability, while the low category consists of the lowest 10% and the conventional category consists of everything in between. Even though there has been a lot of research done on sustainable investing, the results from this research adds a more recent time period that has not been researched yet, furthermore it will take a look at the market timing ability and differing economic cycles. With this an attempt is made to answer the research question "*Does sustainable investing come at the cost of financial performance?*".

To look at the performance of sustainable investing mutual funds, portfolios are made and split into the three categories that are stated above. These three categories will be assessed on their risk-adjusted performance through multiple models, that are widely used throughout the literature. First the hypothesis '*Highly sustainably mutual funds' return and risk (performance do not differ from other types of mutual funds.'* is tested using the CAPM, the Fama-French three-factor model and the Carhart four-factor model. The Jensen's alpha in each of these models is used to assess the performance, leading to mostly insignificant results showing that the three categories do not significantly differ in performance than that of the market benchmark. Regarding the CAPM, the conventional show a positive significant alpha of 0.08, at the 1% level, which implies that the conventional funds show slight outperformance of the market benchmark. Which is in line with various literature discussed above, like Climent and Soriano (2011) and Ibikunle and Steffen (2015). The Sharpe ratios show that the high mutual funds, with a ratio of 0.29, are neither the highest nor lowest category, with the low mutual funds' category with a small lead with a ratio of 0.31. The conventional categories' Sharpe ratio is the lowest at 0.18. This paints a different picture, opposing the CAPM results where the conventional funds showed outperformance. Each category showing a Sharpe ratio less than 1, is in line with Sharpe (1968) where mutual funds show underperformance to the market.

Due to the this and the long time period, further analysis is done; which test the next hypothesis *'Differing economic cycles do not affect the return and risk (performance) of highly sustainable mutual funds.'*. Using the market timing model, the high sustainable mutual funds show significant outperformance to the market benchmark, at the 5% level, in the bull- and bear markets; greater than its conventional or low counterparts. These results are most in line with Ito et al. (2013) where SRI funds show outperformance to conventional funds.

To answer the main question one more hypothesis is tested regarding the market timing ability of each of the categories, '*The return and risk (performance) of highly sustainable mutual funds is not affected by market timing.'*. With Treynor and Mazuy's model added to the Carhart four-factor model the results show

significant outperformance to the market benchmark for the high sustainable mutual funds category, at the 5% level. The low sustainably mutual funds category also shows significant outperformance, but to a lesser extent; with the conventional mutual fund category showing no significant outperformance.

Table 11: Summarized results for Jensen's alphas and Sharpe Ratio

Sustainability rating	*α (CAPM)*	*Three factor α*	*Four factor α*
High	0.02 (0.15)	0.06 (0.40)	0.10 (0.63)
Low	-0.02 (-0.12)	0.12 (1.00)	0.07 (0.50)
Conventional	0.08*** (2.83)	-0.11 (-1.02)	-0.08(-0.70)
High (Bull)	0.90*** (2.81)	0.86*** (2.77)	0.65** (2.02)
Low (Bull)	0.55** (2.14)	0.53** (2.18)	0.25 (0.83)
Conventional (Bull)	0.045** (2.05)	0.43** (2.03)	0.26 (1.16)
High (Bear)	0.45 (1.60)	0.51* (1.86)	0.73** (2.58)
Low (Bear)	0.34 (1.35)	0.42* (1.79)	0.60** (2.40)
Conventional (Bear)	0.29 (1.25)	0.34 (1.55)	0.48** (2.05)
High (dual beta)	0.33* (1.93)	0.35** (2.12)	0.37** (2.15)
Low (dual beta)	0.22 (1.41)	0.29** (2.11)	0.26* (1.71)
Conventional (dual beta)	0.08 (0.66)	0.12 (0.94)	0.12 (0.95)
Fund type:	High	Low	Conventional
Sharpe Ratio:	0.29	0.31	0.18

MPT states that limiting your investment horizons will lead to a less optimal financial performance. In contrary, this research shows that sustainable investing does not necessarily come at a cost of the financial performance. In only one of the models, namely the CAPM, does the high mutual funds underperform the conventional mutual funds, while there is no significant difference between the high and low sustainable mutual funds. Then there are results show that, when accounting for economic cycles or market timing, high sustainable mutual funds tend to outperform the market benchmark and the low sustainable- and conventional mutual funds.

In conclusion, except for the risk-adjusted performance in the CAPM, highly sustainable mutual funds show no underperformance. Further research can investigate this, but for these results a choice must be made when investing in sustainability. If CAPM is the model of choice, then sustainable investing comes

at a slight cost of financial performance. In contrary, if the other models are the main indication, there is no reason to consider that sustainable investing comes at a cost of financial performance.

Some limitations to this research come from the fact that the sustainability rating taken from Morningstar starts from 2016, hence a further analysis of the mutual funds could be done to assess their sustainability before this time period. The long time period could also be divided into shorter time periods to see how the performance changed throughout time; with sustainability getting more attention, this might show a better performance for sustainable mutual funds in the more recent time period. Also, different geographical locations can be analysed like Ito et al. (2013); instead using Jensen's alpha to see if location has an influence on performance.

REFERENCES

Ambec, S., & Lanoie, P. (2008). Does It Pay to be Green? A Systematic Overview. *Academy of Management Perspectives, 22(4).* 45-62.

Blake, C., & Morey, M. (2000). Morningstar Ratings and Mutual Fund Performance. *Journal of Financial and Quantitative Analysis, 35(3),* 451-483.

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

CFA Institute (2015). *Environmental, social, and governance issues in investing.* Retrieved from https://www.cfainstitute.org/-/media/documents/article/position-paper/esg-issues-in-investing-a-guide-for-investment-professionals.ashx

Climent, D. & Soriano F. (2011). Green and good? The investment performance of US environmental mutual funds. *Journal of Business Ethics, 103(2),* 275-287.

Del Guercio, D. & Tkac, P. (2008). Star Power: The effect of Morningstar ratings on mutual fund flows. *Journal of Financial and Quantitative Analysis, 43(4),* 907-936.

Dolvin, S., Fulkerson, J., & Krukover, A. (2018). Do "good guys" finish last? The relationship between Morningstar sustainability ratings and mutual fund performance. *Journal of Investing, 28(2),* 77-91.

Eccles, R., Ioannou, I., & Serafeim, G. (2014). The Impact of Corporate Sustainability on Organizational Processes and Performance. *Management Science, 60(11),* 2835-2857.

EUROSIF A.I.S.B.L. (2008). *European SRI Study*. Retrieved from http://www.eurosif.org/wp-content/uploads/2014/04/eurosif_sristudy_2008_global_01.pdf

EUROSIF A.I.S.B.L. (2016). *European SRI Study*. Retrieved from http://www.eurosif.org/wp-content/uploads/2016/11/SRI-study-2016-HR.pdf

Fama, E. (2014). Two pillars of asset pricing. *American Economic Review, 104(6),* 1467-1485.

Fama, E., & French, K. (1992). The cross-section of expected stock returns. *Journal of Finance, 47(2),* 427-465.

Fama, E., & French, K. (1993). Common risk factors in stock and bond returns. *Journal of Financial Economics, 33(1),* 3-56.

Financial Times Ltd. (2012). *Fund managers focus on ESG factors.* Retrieved from https://www.ft.com/content/d9e42c76-3fda-11e2-9f71-00144feabdc0

Forbes Media LLC. (2018). *The remarkable rise of ESG.* Retrieved from https://www.forbes.com/sites/georgkell/2018/07/11/the-remarkable-rise-of-esg/#e5c7a5416951

Grinblatt, M. & Titman, S. (1992). The persistence of mutual fund performance. *Journal of Finance, 47(5),* 1977-1984.

Guercio, D. & Reuter, J. (2014). Mutual fund performance and the incentive to generate alpha. *Journal of Finance, 69(4),* 1673-1704.

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

Ibikunle, G. & Steffen, T., (2015). European green mutual fund performance: a comparative analysis with their conventional and black peers. *Journal of Business Ethics, 145(2),* 1-19.

Ito, Y., Managi, S., & Matsuda, A. (2013) Performances of socially responsible investment and environmentally friendly funds. *Journal of the Operational Research Society, 64 (1 1),* 1583-1594.

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

Jensen, M. (1967). The performance of mutual funds in the period 1945-1964. *Journal of Finance, 23(2),* 389-416.

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.

Last Word Media UK Limited. (2018). *European ESG flows exceed €32bn in first half of 2018.* Retrieved from https://expertinvestoreurope.com/european-esg-flows-exceed-e32bn-in-first-half-of-2018/

Lintner, J. (1965a). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics 47(1),* 13-37.

Lintner, J. (1965b). Securities prices, risk, and maximal gains from diversification. *Journal of Finance 20(4),* 587-615.

Mallin, C., Saadouni, B., & Briston, R. (1995). The financial performance of ethical investment funds. *Journal of Business Finance & Accounting, 22(4),* 483-496.

Markowitz, H.M. (March 1952). "Portfolio Selection". *Journal of Finance*, 7 (1), 77–91.

Morningstar Inc. (2016). *De Morningstar Sustainability Rating: een nieuw perspectief voor beleggers.* Retrieved from http://www.morningstar.nl/nl/news/148035/de-morningstar-sustainability-rating-een-nieuw-perspectief-voor-beleggers.aspx

Morningstar Inc. (2018). *Morningstar Sustainability Rating Methodology.* Retrieved from https://www.morningstar.com/content/dam/marketing/shared/research/methodology/744156_Morningstar_Sustainability_Rating_for_Funds_Methodology.pdf

Morningstar Inc. (2018). *Does Investing Sustainably Mean Sacrificing Return?* Retrieved from https://www-prd.morningstar.com/content/dam/marketing/shared/pdfs/Research/Does_Investing_Sustainably_Mean_Sacrificing_Returns_March2018.pdf?cid=EMQ_&utm_source=eloqua&utm_medium=email&utm_campaign=&utm_content=17484

Mossin, J. (1966). Equilibrium in a capital market. *Econometrica, 34(4),* 768-783.

Muñoz, F., Vargas, M. & Marco, I. (2014). Environmental Mutual Funds: Financial Performance and Managerial Abilities. *Journal of Business Ethics 124(4),* 551-569.

Nobel Media AB (1990). *The Sveriges Riksbank prize in economic sciences in memory of Alfred Nobel 1990*. Retrieved from https://www.nobelprize.org/prizes/economic-sciences/1990/press-release/

Paramita V. (2015). Testing Treynor-Mazuy Conditional Model in Bull and Bear Market. *Review of Integrative Business and Economics Research, 4(2),* 208-219.

Renneboog, L., Ter Horst, J., & Zhang, C. (2008a). Socially Responsible Investments: Institutional Aspects, Performance, and Investor Behavior. *Journal of Banking & Finance, 32(9),*1723-1742.

Renneboog, L., Ter Horst, J., & Zhang, C. (2008b). The price of ethics and stakeholder governance: The performance of socially responsible mutual funds. *Journal of Corporate Finance, 14(3),* 302-322.

Sharpe, W. (1964). Capital asset prices: a theory of market equilibrium under conditions of risk. *Journal of Finance, 19(3),* 425-442.

Sharpe, W. (1966). Mutual fund performance. *Journal of Business, 39(2),* 119-138.

Treynor, J. (1961). Market value, time, and risk. Unpublished manuscript.

Treynor, J. (1962). Toward a Theory of Market Value of Risky Assets. Unpublished Manuscript.

UNEP Finance Initiative (2005). *Who cares wins, Connecting financial markets to a changing world.* Retrieved from https://www.unepfi.org/fileadmin/events/2004/stocks/who_cares_wins_global_compact_2004.pdf

United Nations (n.d.). *Sustainable development goals.* Retrieved from https://sustainabledevelopment.un.org/?menu=1300

US SIF: The Forum for Sustainable and Responsible Investment. (2016). *The impact of sustainable and responsible investment.* Retrieved from https://www.ussif.org/files/Publications/USSIF_ImpactofSRI_FINAL.pdf

US SIF: The Forum for Sustainable and Responsible Investment. (2019). *SRI Basics*. Retrieved from https://www.ussif.org/sribasics

APPENDIX

Table A.1: Empirical results for the market timing model Regressions using the *Kenneth R. French data library-sourced* global market factor

This table reports the regression results, with robust standard errors, for the market timing model using a global market factor. The global factor portfolios from the *Kenneth R. French data library* are used as factors to measure the risk-adjusted returns of each of the portfolio categories. α measures of the risk-adjusted abnormal return relative to the market proxy in Bull and Bear markets. Bull- and bear markets are defined as a time period wherein the market shows, respectively, positive- or negative market returns. β_{MKT} measures the effects of the MKT factor. The t-statistics are depicted in parentheses. *, ** and *** correspond to statistical significance at 10%, 5% and 1% levels respectively. The time period is July 1990 to December 2018.

Sustainability rating	α	β_{MKT}	*Adj.* R^2
High (Bull)	0.90*** (2.81)	0.48*** (4.89)	0.14
Low (Bull)	0.55** (2.14)	0.76*** (11.54)	0.33
Conventional (Bull)	0.45** (2.05)	0.65*** (9.48)	0.39
High (Bear)	0.45 (1.60)	0.89*** (14.21)	0.55
Low (Bear)	0.34 (1.35)	1.03*** (17.25)	0.66
Conventional (Bear)	0.29 (1.25)	0.97*** (19.89)	0.71

Table A.2: Empirical results for the market timing model Regressions using the *Kenneth R. French data library-sourced* global market factor

This table reports the regression results, with robust standard errors, for the market timing model using a global market factor. The global factor portfolios from the *Kenneth R. French data library* are used as factors to measure the risk-adjusted returns of each of the portfolio categories. α measures of the risk-adjusted abnormal return relative to the market proxy in Bull and Bear markets. Bull- and bear markets are defined as a time period wherein the market shows, respectively, positive- or negative market returns. β_{MKT}, β_{SMB} and β_{HML} measure the effects of the MKT, SMB and HML factors. SMB corresponds to the return spread between a small cap portfolio and a large cap portfolio and HML is the difference in return between a value stock portfolio and a growth stock. The t-statistics are depicted in parentheses. *, ** and *** correspond to statistical significance at 10%, 5% and 1% levels respectively. The time period is July 1990 to December 2018.

Sustainability rating	α	β_{MKT}	β_{SMB}	β_{HML}	*Adj.* R^2
High (Bull)	0.86*** (2.77)	0.49*** (5.32)	0.24* (1.81)	-0.09 (-0.88)	0.17
Low (Bull)	0.53** (2.18)	0.77*** (11.65)	0.26 (1.37)	-0.30** (-2.04)	0.38
Conv. (Bull)	0.43** (2.03)	0.65*** (9.97)	0.16 (1.51)	-0.12 (-1.30)	0.41
High (Bear)	0.51* (1.86)	0.87*** (13.75)	0.08 (0.64)	-0.18** (-2.19)	0.57
Low (Bear)	0.42* (1.79)	0.98*** (19.17)	0.27** (2.29)	-0.32*** (-3.92)	0.72
Conv. (Bear)	0.34 (1.55)	0.95*** (19.78)	0.09 (0.87)	-0.17** (-2.56)	0.72

Table A.3: Empirical results for the dual beta model Regressions using the *Kenneth R. French data library-sourced* global market factor

This table reports the regression results, with robust standard errors, for the Treynor and Mazuy model using a global market factor. The global factor portfolios from the *Kenneth R. French data library* are used as factors to measure the risk-adjusted returns of each of the portfolio categories. α measures of the risk-adjusted abnormal return relative to the market. β_{MKT} measures the effects of the MKT factor. The t-statistics are depicted in parentheses. *, ** and *** correspond to statistical significance at 10%, 5% and 1% levels respectively. The time period is July 1990 to December 2018.

Sustainability rating	α	β	B_5	*Adj.* R^2
High	0.33* (1.93)	0.71*** (19.66)	-0.02*** (-3.51)	0.58
Low	0.22 (1.41)	0.90*** (26.20)	-0.01** (-2.35)	0.71
Conventional	0.08 (0.66)	0.81*** (26.31)	-0.01*** (-2.77)	0.77

Table A.4: Empirical results for the dual beta model Regressions using the *Kenneth R. French data library-sourced* global market factor

This table reports the regression results, with robust standard errors, for the Treynor and Mazuy model using a global market factor. The global factor portfolios from the *Kenneth R. French data library* are used as factors to measure the risk-adjusted returns of each of the portfolio categories. α measures of the risk-adjusted abnormal return relative to the market. β_{MKT}, β_{SMB} and β_{HML} measure the effects of the MKT, SMB and HML factors. SMB corresponds to the return spread between a small cap portfolio and a large cap portfolio and HML is the difference in return between a value stock portfolio and a growth stock. The t-statistics are depicted in parentheses. *, ** and *** correspond to statistical significance at 10%, 5% and 1% levels respectively. The time period is July 1990 to December 2018.

Sustainability rating	α	β_{MKT}	β_{SMB}	β_{HML}	B_5	*Adj.* R^2
High	0.35** (2.12)	0.70*** (17.11)	0.18* (1.95)	-0.13** (-1.98)	-0.02*** (-2.85)	0.60
Low	0.29** (2.11)	0.88*** (26.76)	0.26** (2.01)	-0.32*** (-3.35)	-0.01** (-2.17)	0.75
Conventional	0.12 (0.94)	0.80*** (26.77)	0.13* (1.76)	-0.14** (-2.52)	-0.01*** (-2.72)	0.78