Universitetet i Stavanger

**UIS BUSINESS SCHOOL**

# MASTER'S THESIS

| STUDY PROGRAM: | THESIS IS WRITTEN IN THE FOLLOWING SPECIALIZATION/SUBJECT: |
|---|---|
| Master of Science: Business and Administration | Applied Finance |

TITLE:

Do SRI investors consider the ESG effects in their investments?


| AUTHOR(S) | | SUPERVISOR: |
|---|---|---|
| Candidate number: | Name: | Klaus Mohn |
| 4017 ………………… | Sindre Johansen ……………………………………. | |
| 4058 ………………… | Eirik I. Åse ……………………………………. | |

# Abstract

In this master thesis we explore the relationship between morally driven companies and their financial returns. To assess the morality of the company, we have used scores based on a company's environmental, social and governmental performance. The scores are based on a trusted $3^{rd}$ party rating agency. By adding an ESG-score term to the Single-index model and the Fama-French-Carhart model, we were able to empirically assess the effect of each point of ESG score against the return of a given company.

Based on 10 years of historical data we were able to find both statistical and economically significance that each point of ESG turned out to reduce the expected return annually with 0,147 %.  Moving from the 25 % highest rated to the 25 % lowest rated in terms of ESG (in our sample) would yield an annually increased return of 3,10 %. This numerical figure was the same for the single-index model and the Fama-French-Carhart model.

We also test for country differences and sector differences where we find statistical significance differences in return for Italy and France and the Energy sector. These differences are also correlated with the return in ESG score.

Furthermore, we reduced our sample to check if there were differences when using 5 years, 2 years and 1 year of data. Results then showed to be inconsistent and we could not find statistical significance for these estimated time periods.

# Table of contents

## Figures



## Tables

# 1.0 Introduction

## 1.1 Topic question

*"Do SRI investors consider the ESG effects in their investments?"*

In this thesis we want to explore if investors driven by the concept of socially responsible investing (SRI) consider the level of the environmental, social and/or governance (ESG) that an investment protrudes. SRI investors are driven by investments that contribute to a public good, but the question remains if SRI is merely another requirement to check of from a fixed list of demands for the investor. This topic question has a two-fold purpose. It measures if SRI investors are willing to sacrifice financial return to contribute to a public good, and if so, how much return the investor is willing to sacrifice for a given amount of public good.

## 1.2 Defining Socially responsible investing and ESG

Socially responsible investing is hard to define. The ambiguous definition of SRI may cause inquiries to the comparability of different papers on the topic. As a way of measuring SRI, the "ESG criteria" can be used. For an investment to fulfil the ESG criteria, it must provide some sort of value regarding Environmental, Social or Governmental factors. E.G; Investments that slow climate change development, promotes anti-corruption policies or increases board member diversity could be regarded as socially responsible investments. It is also important to note these are investments and not donations; the investor is seeking a financial return coherent with the size of his investment.

The ESG criteria provide clarity to the issue regarding the ambiguousness of the definition of SRI. ESG is not directly compared with SRI, neither is the term impact investing. ESG is measured based on a scoreboard with guidelines regarding the value of ESG. It is easier for practitioners to follow but subject to abuse as the scores are easy to manipulate. SRI is more of an ethical guideline where investors should allocate their assets to investments that provide benefit to the society while shy away from investments that are harmful to the society. As SRI is hard to rank in a linear analysis, we will, due to their similarities, group the definitions of ESG and SRI together.

ESG is a generic term which is used by investors to evaluate corporate behaviour. How companies are defined in terms of ESG framework is in theory different from every individual investor or company. To align the individual investors or company, using an acknowledged rating agency may be enough for analysis. A potential pitfall in using the ESG criteria as a proxy for SRI is that it does not include the act of negative screening of companies that the SRI framework enforces.

### 1.3 Background and motivation

The idea behind investments that generate value beyond financial returns, socially responsible investing, has been around for decades. SRI, familiar with corporate social responsibility (CSR), has gained traction in modern times and have transformed from negatively screening of "sin" companies (tobacco, slave-trade, alcohol) to concern more about sustainable investing and ethical business behaviour.

A landmark study, "Who cares wins" initiated by the The Global Compact, International Finance Corporation and Federal department of foreign affairs Switzerland in 2005 coined the term ESG to help better integrate such issues in analysis, asset management and securities brokerage. The ESG definitions are well explained and its comparability with other companies give investors more well-defined information.

The hype towards ESG and SRI has made bold claims towards the investments financial return. To showcase this hype, Forbes (2017) wrote an article advocating for positive correlation between ESG scores and financial return. While such claims are not necessarily wrong, they contradict financial logic and portfolio theory where investors seek to maximize return with minimal risk. If the optimal combination of risk and return does not provide the qualities of ESG as an SRI investor wished it had possessed, it should not be able to exceed traditional investments in terms of risk and return.

If the claims towards increased financial return on highly rated ESG investments are true, investments such as these should see an influx of not only morally concerned investors but also traditional investors who want to reap the rewards of increased financial return. If traditional financial and microeconomic logic is true, investments who are SRI-concerned should provide less value in terms of financial return. A traditional investor could then create a trading strategy which would yield higher financial return by investing in lower rated ESG companies. The results from this thesis could have major implications with significant meaning and this is our key motivation behind writing a thesis on this topic.

## 1.4 Structure

In the second and third chapter we investigate relevant research to our topic question and methodology. We explore several established financial theories and models and attempt to grasp the development from the Capital Allocation Line by Markowitz in 1959 all the way to the 4-factor model by Fama-French-Carhart in 1997. Furthermore, we investigate different research conducted within the area of SRI. By using the research chapter as a base, we explore each separate component of several financial models. The third chapter also includes a review of how to measure the ESG-scores, an explanation of the concept of utility, a discussion regarding the proper statistical procedure and in the end a construction of our regression estimation.

The fourth chapter showcase the various data we have used in our estimation, and the sources of this data. Most of the data have been retrieved by using the Thomson Reuter Datastream software.

The fifth chapter present our regression estimation and the regression results. On the background of these results, we interpret the results and understand them in connection to our topic question. This further leads to our conclusions of the paper in chapter six and suggestions for further research on this topic in chapter seven.

## 2.0 Research

This chapter discusses related research that has been conducted prior to this paper. First, we cover some of the earlier findings on theoretical finance and modern portfolio theory. The research on modern portfolio theory dates back all the way to Markowitz in 1959. In the time after, several models were created with Markowitz's research as a foundation. However, these models proved to be no empirical success, and did not hold true in real markets. Consequently, this further lead to more research and more detailed models that tried to explain and solve shortcomings from the prior ones. Furthermore, we discuss the previous research which combine portfolio theory to socially responsible investing and observe the strength and weaknesses of the different methods used. Finally, in the end we contract the issue down to regional differences and discuss aspects around this.

Markowitz (1959) introduced a model to optimal portfolio choice. The model assumes a risk averse investor that only care about the mean and variance of their one-period investment return. This introduced the term "Mean-variance-efficient portfolios" which is based upon investors choosing portfolios that minimize the variance of portfolio return, given a certain level of expected return and maximize expected return, given a certain level of variance. With the assumption that investors only care about systematic risk, it further led to the Capital Allocation Line (CAL), which show an investors optimal portfolio based on that investors level of risk. This has later been called the birth of modern portfolio theory. Markowitz's initial research has been one of the staples of theoretical finance and has led to a large volume of research expanding on this topic.

Several people investigated Markowitz's theory, and tried to enhance it: Jack Treynor (1961, 1962), William F. Sharpe (1964), John Lintner (1965) and Jan Mossin (1966). They added two key assumptions to the Markowitz model to be able to identify a portfolio that must be mean-variance efficient.

The first added assumption was that there must be complete agreement between all investors regarding the market clearing asset prices at t-1 to t and the joint distribution of asset returns in the same period. It is also assumed that this distribution is the true one.

The second added assumption was that all investors can borrow and lend at the risk-free rate and do not depend on the amount that is borrowed or lent. This led to all efficient portfolios being different combinations of the risk-free asset and a single risky tangency portfolio. This combination would vary depending on the investors level of risk.

By adding those two assumptions and using Markowitz as the foundation, the Capital Asset Pricing Model (CAPM) was developed. The following relationship was presented:

**(1)** $$R_i = R_f + \beta_{\mathrm{i}}(R_m - R_f)$$

CAPM standardizes the trade-off between risk and return in an unobservable market-portfolio. The expected return of any asset is a function of the risk-free rate plus a risk premium multiplied by the asset's market beta. The results in the model explains what rate of return investors should receive for investing in an asset at a given level of systematic risk (beta).

The CAPM model is based on unrealistic assumptions, and hence would be a simplicity of the reality. However, by introducing the model, it has undertaken extensive testing and paved way for a lot of new research and models.

One of the first to explore the CAPM further, was Jensen (1968). He noticed that the suggested relationship between the expected return and the market beta implied a time series regression test. The data consisted of 115 open end mutual funds whose net assets and dividend information were retrieved from Wiesenbergs Investment Companies for a ten-year period between 1955-1964.

The CAPM model suggested that the expected value of an asset's excess return could be completely explained by its expected CAPM risk premium. Jensen suggested that this may not be the case, and that there could be excess returns for assets. He created a new empirical model:

**(2)** $$\alpha = R_p - (R_f + \left(R_m - R_f\right)\beta_i)$$

The model explains excess returns (returns on security subtracted by the risk-free rate) in terms of a constant and the relative return to the market. This constant was later dubbed "Jensen's Alpha".

If the first CAPM formula was correct, that would imply that the Alpha value would be zero for every asset. This however was not the case in Jensens' analysis. He received alpha values on many of the funds that was below or above zero. However, on average the mutual funds were not able to do significantly better than what one could expect from mere random chance. This means that an excess return at one time, does not imply excess return at a later stage.

There has also been critique towards the CAPM. Roll (1977) argued that the CAPM never had been tested properly and never would. He especially pointed out the fact that the true market portfolio is non-observable because the true market portfolio would consist of all individual assets. One would not know which assets that could be excluded from the true market portfolio, e.g. human capital. As a result of not being able to observe the true portfolio, all the testing would be based upon proxies of the market portfolio. This has later been dubbed as "Roll's critique."

Banz (1981) investigated the relationship between the return and the market capitalization of stocks. He applied a market size term to the traditional CAPM model. He used a sample of common stocks quoted on the NYSE for at least five years in the period 1926 and 1975 collected from the Center for research and security prices (CRSP). By applying both OLS and GLS regression, he found that the additional term for market size reported negative values. This meant that he uncovered that average returns on small capped stocks were higher than predicted by the CAPM. This has later been dubbed as the “small-cap bias.”

Ross (1976) introduced a proposed alternative to the mean-variance CAPM. He introduced the Arbitrage Pricing Theory (APT) which stated that the return on the assets could be explained by using a linear relationship between the assets’ expected return and macroeconomic variables. APT assumes that markets are not always efficient, and sometimes would misprice securities as an effect of this. This could lead to potential profit for arbitrageurs that takes advantage of the mispricing. The following relation was introduced:

**(3)** $$R_i = E(R_i) + B_i F_i + \mathcal{E}_i$$

The idea is that the return of the asset can be predicted by using the linear relationship between the expected return of the asset and several macroeconomic variables that capture systematic risk. One of the problems, however, has been to identify the different macroeconomic variables.

Fama and French (1993) expanded upon the concept of APT and CAPM. They proposed two market-wide variables which should be more capable of predicting a securities’ return rather than using the basic CAPM formula. The two factors that were introduced was the small-minus-big factor (SMB) and the high-minus-low factor (HML). They are meant to capture the size of the company (in terms of market value) and to mimic the returns related to book-to-market equity. The procedure to calculate the factors are done by dividing stocks into size ranked deciles. Then each size is sub-divided into pre-ranked beta deciles and grouped into a total of six portfolios: “Small Value”, “Small Neutral”, “Small Growth”, “Big Value”, “Big Neutral” and “Big Growth”.

The SMB factor is constructed as:

**(4)** $$SMB = \frac{1}{3}(Small\ Value + Small\ Neutral + Small\ growth) - \frac{1}{3}\ (Big\ Value + Big\ Neutral + Big\ Growth)$$

The concept is that if the SMB factor is positive, it is possible to get abnormal returns by investing only in small companies. This procedure is to account for the small cap bias that Banz (1981) discovered.

The HML factor is constructed as:

**(5)** $$HML = \frac{1}{2}(Small\ Value + Big\ Value) - \frac{1}{2}(Small\ Growth + Big\ Growth)$$

A positive HML score indicates that investing in companies with high book values would yield abnormal returns.

Expanding even further, Carhart (1997) explored the common factors that drives mutual funds risk-adjusted returns. He employed both the CAPM and the Fama and French 3 factor model. He added another factor, the momentum factor (MOM). It is created by using six value-weight portfolios formed on size and prior returns.[1] This created the Fama French Carthart 4-factor model.

The MOM factor is constructed as:

**(6)** $$MOM = \frac{1}{2}(Small\ High + Big\ High) - \frac{1}{2}(Small\ Low + Big\ Low)$$

[1] See the Fama-French official website for in depth explanations of the different portfolios.

A positive MOM factor would indicate that an investor could achieve abnormal returns by investing in stocks that recently have had a surge in their stock price.

Carhart analysed the MOM factor by using data from 1892 funds on the S&P which included 582 dead funds to account for survivorship bias. He used a value-weighted CRSP return index which includes all stocks on NYSE, Amex and Nasdaq. An OLS regression was run on the single-index model, the Fama-French three factors, and the Carhart four-factor model. He further added an expense factor and a turnover ratio which accounted for the cost of asset management.

He found that turnover reduces performance for about 95 basis points for every buy and sell transaction. He also found that load-funds underperform no-load funds by approximately 80 basis points per year. His analysis further concluded that active investment management and manager skill have little to do with superior fund performance.

The research covered up to this point has included some of the staples from theoretical finance. Modern portfolio theory and especially the CAPM has both its strength and flaws, and research have been conducted to find adjustments or/and alternatives. The Fama-French factors were ground-breaking in terms of empirical finance as the factors has proven robust in use and consistent over time. We will now narrow down the chapter to research done on SRI perspectives and the conflicting results that different researchers have gathered.

Elton, Gruber, Das & Hlavaka (1993) investigate the informational efficiency of mutual fund performance. They test performance in terms of alpha returns using both a single-index model and a three-index model. They collect data on mutual funds and bonds from the period 1945-1984 from S&P.  S&P were used as benchmark for larger cap stocks and for smaller stocks, CRSP return indexes were used. For bonds, both the Shearson-Lehman index and conventional bond funds were used as benchmark. They use OLS regression on the funds by using both a single-index regression on the excess returns of the S&P and a three-index regression on the excess returns of S&P, excess return on the small-stock index orthogonalized on the S&P and an 80:20 bond index orthogonalized on both the S&P and small stock index returns. They find that active fund managers underperform passive portfolios and that fund with higher fees and turnovers underperform funds with lower fees and turnover; implying a cost in active fund management whose benefit does not make up for it. Considering that SRI funds need to be actively managed (screening and governance to keep up an ethical profile), this could become a problem in the performance when measuring funds.

Lee, Humphrey, Jacquelyn, Karen & Jason (2010) explores the socially responsible funds' performance when imposing non-financial screens. They hypothesize that non-financial screens reduce investment opportunities which will reduce diversification efficiencies and thus affect investment performance. They test for several factors including the Morningstar Squared, Carhart model and CAPM model to estimate the effect of non-financial screens. Their sample consist of 61 US equity SRI funds which are reported from the Social Investment forum. Index data was sourced from the Morningstar Direct Databases and the factors to the Carhart-performance model was obtained from the Kenneth French Data Library. The performance measures were set as the OLS dependent variable and was measured against the screening intensity for the funds, the age of the fund, the size of the fund and whether the fund was an institutional fund, if the fund can vote in proxy policy and if the fund has had other flow than equity during the period. They find no effect on unadjusted (raw) risk. However, when using the Carhart performance model they find a 70-basis point reduction in alpha-returns which are statistically significant. They also find funds with screening to inherit lower systematic risk due to the selection of lower beta-stocks. This could explain that the socially responsible funds pick "safer" stocks due to their limitations of diversification in the investment universe.

A regional study conducted in Australia by Bauer, Otten, Rad (2006) explores the performance and investment style of retail ethical funds. They apply the Carhart four-factor model in their measure of fund performance, alpha. Their sample consist of pure retail equity funds in a total of 25 ethical funds and 281 conventional funds. For the benchmark index they use a proxy supplied from Worldscope which covers up to 98 % of the total market capitalization. When estimating the performance, they undertake an OLS regression which adjust for a home-bias and a sensitiveness to time. They find that ethical portfolios underperform in the beginning of the 1990's but, contrary to Lee, Humphrey, Jacquelyn, Karen, Johnson (2010), find that the ethical funds quickly catch up to match conventional funds in term of performance during the 1996-2003 period. They conclude that investors do not face a financial penalty by selecting SRI funds.

Schrôder (2004) studies the performance of socially responsible equity investment fund's in the United States, Germany and Switzerland and of SRI equity indices such as Domini 400-index. He estimates the performance as Jensen's alpha by using a special case of the single-index model where he divides the market portfolio in a separate beta for blue chip stocks and small-cap stocks to address the problem regarding small company bias.  Financial return for 30 funds from the US, 16 funds from Germany and Switzerland and 10 SRI indices are used as a sample. OLS regression was carried out by using three different versions of the special case single-index model which accounts for the timing of the market and changing market conditions. They find that some of the SRI funds underperform in measure of Jensen's alpha and the result is statistically significant, while most of the SRI indices have an insignificant, but positive alpha. Regardless, they conclude that on average, SRI investors do not receive a financial penalty by investing in SRI assets.

Another study, conducted by Bello (2005), takes a different approach by examining the differences in characteristics of asset held, portfolio diversification and variable effects of diversification on portfolio performance on socially responsible funds vis-à-vis conventional funds.  He measures the portfolio performance in terms of both variance and returns. The data is sampled from 42 responsible funds (including 3 dead funds) and 84 conventional funds from the Morningstar March 2001 principa pro database. He also finds corresponding returns on the S&P 500, monthly return data from the DSI 400 and three-month treasury bills. He uses three alternate measures for investment performance; the single index model, the Sharpe information ratio and the excess standard deviation adjusted return. He finds that the conventional and SRI funds do not differ in term of asset characterises, degree of portfolio diversification or long run investment performance. He further finds that both groups of funds have significant extra market covariation, indicating that both groups of funds are undiversified and underperforms relative to the benchmark index. Bauer, Derwall, Otten (2006) further adds that differences between SRI funds and conventional funds not necessarily are transparent. They find that the SRI funds correlate more with conventional market indices rather than with ethical indices.

Timothy M. Doyle (2018) wrote a report on “Ratings that don’t rate – The subjective world of ESG ratings agencies”, an ESG criticism on behalf of the American Council for Capital Formation. As a warning; The ACCF is a think-tank founded in 1975 by Charls Walker. The foundation itself describes itself as nonpartisan, but journalists generally describe them as free market or pro-business. Walker himself served under republican presidents, and hence one must beware that his foundation might publish articles that waves in their favour.

The report highlights three possible biases with the ESG-rating system. He argues that ratings agencies apply a one-size-fits-all approach which has provided benefits for larger companies. According to an analysis of 4150 Sustainalytics[2] ESG ratings show that larger companies tend to obtain better ratings. It is theorized that larger companies might be in a better financial shape and therefore are able to invest more in measures that improve their ESG profile. Many of those companies also have resources to publish comprehensive annual ESG disclosures. As a result, small size and mid-sized companies are at a competitive disadvantage. Doyle looks at several small pharmaceutical companies with lower than average ratings, but argues that their efforts, even though they are aligned with ESG issues, are not properly reflected in the rating. This kind of bias could make smaller companies less appealing for SRI investors.

The second bias is based upon the geographical location of the different companies, and the differences in the reporting requirements. Doyle points out that disclosure requirements vary significantly by country and region, and that this is the primary source of information the rating providers use. This suggests that countries with high disclosure requirements also would have the best base to get higher ESG ratings. To support this claim, he has found that the EU in Europe requires companies with 500

[2] Sustainalytics is a company that rates the sustainability of listed companies based on their ESG performance.

or more employees to publish a “non-financial-statement” and additional disclosures around diversity policy, while North America has no such requirements. He also compares BMW Group (European) and Tesla (American), and finds that BMW Group, despite a lot of controversy around ESG matters, still manages to outperform Tesla. BMW is in the $93^{rd}$ percentile, while Tesla is only in the $38^{th}$ percentile. This suggests a geographical bias as Tesla is known to be a world leader in technology to reduce carbon emissions from automobiles but holds only an average score.

The third and final bias is about that rating agencies claim to normalize ratings by industry but fail to factor in company-specific risks. Doyle claims that this could result in a biased rating for a company and mislead investors. He exemplifies by comparing large dividend-focused companies within utilities with high average ratings against less mature companies such as pharmaceuticals with focus on reinvesting and R&D and lower average ratings. He argues that the rating system must be better tailored for company-specific risks. Companies within the same industry do not necessarily have all the same risks, and hence should be viewed individually rather than collectively in the industry.

A study directed towards the US finds Tamimi & Sebastianelli (2017) exploring the state of the S&P 500 companies’ transparencies towards ESG score and ESG ratings. They conduct a nonparametric procedure analysis at a granular level. Their data consist of 347 companies that were collected from Bloomberg using the financial analysis environmental, social and governance function for the companies compromising the S&P 500. There is a mix between quantitative and qualitative data. The Kolmorgov-Smirnov goodness of fit test was used to establish market deviations from normality. They find that most S&P 500 companies are transparent regarding governance disclosures but have significant deficiencies regarding the closing of information related to both environmental and social issues. This asymmetrical information regarding the reporting standards in the US may cause ESG investors to not perceive the companies as reliable SRI investments.

Rennebog, Ter Horst, Zhang (2007) estimates the price of ethics by studying the risk-return relation in SRI funds on different regions. They measure performance in alpha and use the Carhart-four factor model to determine superior performance. The data collection is from SRI funds domiciled in Europe, the US and selected countries outside Europe. As a counterpart, conventional fund performance from US and UK have also been collected as a reference group. OLS regression is carried out on an extension of the Carhart-four factor model which includes an ethical factor to determine its possible effect on the stock price. They find that SRI investors explicitly deviate from the economically rational goal for wealth maximization as the average risk adjusted returns in several countries are lower than -5 % per annum. They also find that European passive portfolios which does not include “sin” stocks underperform the benchmark factor by 4.5 % per annum, while they find insignificant relationships in the US. The reporting standards of ESG are stricter in the EU which could mean that SRI investors have a bigger belief in the SRI investments of the EU. Another study by Auer & Schuhmacher (2016) finds the same results regarding European SRI fund performance. They also however find that the scale of the premium is dependent on which ESG factor is prevalent, and that the differences are of significant magnitude depending on the factor.

There are conflicting results regarding the effect that SRI has on financial return. Several methods have been used, but mostly used have been the Carhart four factor model and the CAPM model. The research reveals findings that many SRI funds with on average lower beta stocks, usually have “smaller” companies in the portfolio and tend to converge in terms of ESG rating with their conventional non-ESG funds. This can make comparing the value of ESG of funds biased.

A method to analyse ESG is to use 3rd party rating agencies. Large companies tend to get better rating, there are different geographical differences in terms of ESG-score rating and company specific risk are usually not accounted for when assessing the ESG-score of a company. A study also finds that in the US where most studies on SRI are done, there exists deficiencies when closing information regarding environmental and social issues. This may have led to research in the US not finding any effect of ESG-performance on financial return. In relation, the EU standard is stricter and requires companies to be more transparent on the matter. Two studies that include the European sector, Rennebog, Ter Horst, Zhang (2007) and Auer, Schuhmacher (2016) finds that higher ESG-rated stocks in the European market perform worse than lower ESG-rated stocks, and the results are statistically significant.

## 3.0 Methodology

This chapter presents a review on the foundation of our empirical research. We introduce the aspect of utility and the build-up to the Capital Allocation Line. Further we explain the different components in the Capital Asset Pricing Model and investigate the different approaches and models that has been derived subsequently from the CAPM. By investigating each approach in detail and comparing them, we attempt to determine which model is the most fitting for our research topic.

We then further discuss how we should go about ranking the ESG scores and where to collect them from. We revisit the problems researchers have encountered using ESG scores and present possible solutions to the issues.

Finally, using this chapter as a theoretical foundation, we build our regression model which we will use for our analysis.

### 3.1 Theoretical Background

#### 3.1.1 Expected return

The expected return is the predicted future return of an asset. Using historical data, each observation is treated as an equally likely scenario and the expected return is estimated as the arithmetic average of each observation rates of return:

**(7)** $$E(r) = \frac{1}{n}\sum_{s=1}^{N} r(s)$$

Where $E(r)$ is the expected arithmetic return and $r(s)$ is the return for each observation.

If the time series represented the true underlying probability distribution, then expected return from a historical period would provide a relevant forecast of the investments expected future return.

### 3.1.2 Risk

Risk is in this case measured in standard deviation of a security's return. More risk equals more uncertainty about the movements of the stock price. Risk is often divided in two categories, independent and common risk. Independent risk is risk that does not correlate with other risk factors while common risks are correlated with some or more risk factors.

In established literature, independent risk is referred to as unsystematic risk and is often disregarded in portfolio practice. Adding enough stocks in a portfolio will average out the unsystematic risk and make the portfolio well diversified.

Common risk is related to systematic risk that cannot be diversified away. This risk is relevant for investors and is taken into consideration when planning investment decisions. The systematic risks tend to fluctuate with market movements and macroeconomic conditions.

### 3.1.3 Concept of utility

To analyse an individual choice, we need a presumption on his preference. A starting point is to assume that the individual will behave rational so that he will always do what benefits him the most. As an expression for this preference, "utility" is a creation which helps us rank choices. Utility can be regarded as something to be desired. This desire is symmetric, when something is not desired one can say that it provides disutility.

In microeconomic literature, utility is defined under a set of axioms to define its concept. It demands *completeness* in the sense that one can always rank choices as either better, worse or equal. This assumes that individuals are not paralyzed by indecision. Another axiom of utility, *transitivity* states that choices should be ranked in relation to each other. To put it in other words, if individual prefer choice B over choice C, and prefer choice A over choice B, then they should also prefer choice A over choice C. The last axiom, *Continuity,* says that if an individual prefers A over B, then choices that are closely related to A should also be preferred over choice B.

The problem with utility is that it is impossible to measure with 100 % certainty. One can observe choices between individuals and rank them according to their choices, but these choices can be far from consistent. The reason for recording utility is to find and estimate for an individual's choice at a later state. When all individuals have different preferences one can see that predictions can get noisy and blurry. It also does not help that individuals also change preferences over time. Utility also find itself to have diminishing returns, which means that the more you have of an asset the less you desire it.

Utility, however, is an easy way to define and charter choices. While the assessment of such practice is hard, the theoretical part is easy to explain. Some areas are easier to use utility than others. For investors, utility can be a tool to assess the choices between risky and less risky assets. Investing in assets usually yield few *significant* variables which affects utility or disutility and can thus prove as a useful tool given the right usage. In modern portfolio theory, they usually narrow down to two variables; return which yields utility and risk, which yields disutility.

### 3.1.4 Efficient markets

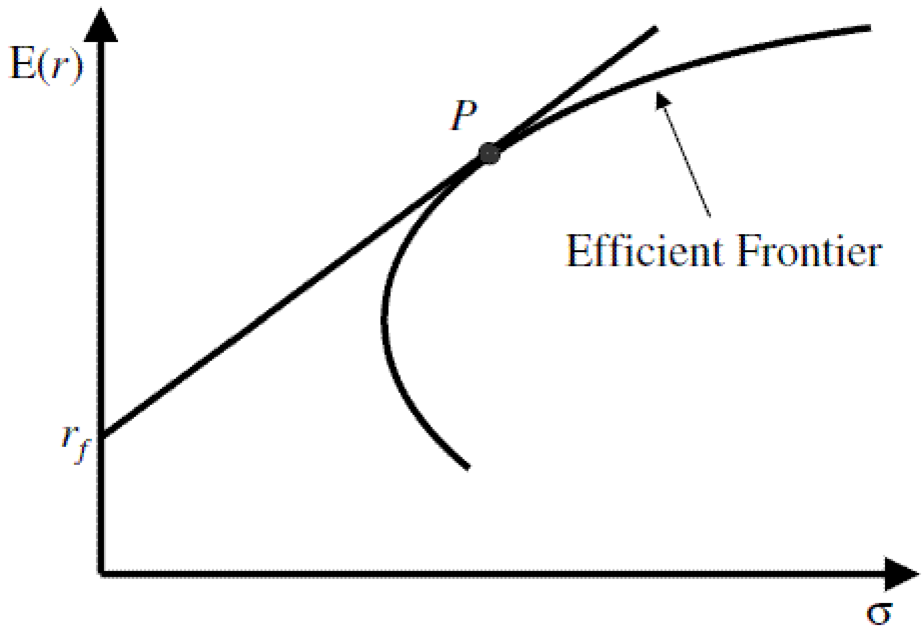
As explicitly stated in the last subsection, financial theory assumes investors to behave rationally. This implicitly assumes that markets will behave rationally; all security prices are priced at its expected value. A security's price would thus translate to the average investor's belief of the expected value of the security. This assumption allows for deviations in stock price due to surprises or unexpected shocks which change the market conditions.

Not all investors have access to the same information, some investors have inside information and some information may be misleading or overstated; which can lead to deviation of its true value. In this analysis we assume that investors behave rational, so markets are efficient in terms of utility maximization.

### 3.1.5 Capital Allocation Line

An investor is seeking returns and is reluctant to risk. As established, an investor can disregard unsystematic risk with proper diversification and will thus only care about the systematic risk. In a world where investors only care about the risk and return, all investors will have the same optimal portfolio regardless of their taste of risk. This is because the optimal portfolio will give the best ratio of risk and return, and by borrowing or lending at the risk-free rate, one can elevate or de-elevate the amount of return one gets, and the risk one bears.
Assuming short sales are possible for all stocks in a portfolio, this situation was illustrated by Markowitz (1959):

*Figure 1: The traditional Capital Allocation Line*




Source: Analystnotes.com

The tangency point between the efficiency frontier and the capital allocation line is the attribute to risk and reward that the optimal portfolio inherits.

### 3.1.6 Utility Curve

While all investors in a “risk-return world” will derive the most utility by holding the same portfolio, investors have different risk appetite. Risk-appetite is hard to quantify but a proposed utility curve for an individual investor for determining the optimal level of return/risk is:

**(8)** $$U = E(r) - 0{,}5A\sigma^2$$

Where $\sigma^2$ is the variance of the returns and $A$ is the level of risk aversion per individual investor.

When $A > 0$, investors are considered risk averse. Investors that are risk averse gets disutility by increased level of risks. Investors are risk averse and the logical foundation to the capital allocation line is built upon this assumption. If most investors were not risk-averse, there would not exist a risk-return trade-off as illustrated in financial literature.

When $A = 0$, investors are considered risk neutral. Investors that are risk neutral judge investments solely on their expected rates of return and disregards the concept of risk. A risk-neutral investor’s required rate of return for any risky project could theoretically be as small as the risk-free rate, given that all other investors are risk-neutral.

When $A < 0$, investors are considered risk-lovers. Risk-loving investors exhibit “gambling-like” behaviour in which they derive utility from increased risk in projects.

### 3.1.7 Risk free rate

The risk-free rate is the rate which investors can borrow and save at. It is also considered the rate that risk-free investments will return. The risk-free rate is usually estimated by using the yields of assets that are usually considered risk free; common practice is to use government bonds. Both the length and rate of return on government bonds differ, but it is agreed that the horizon of the government bonds should match the investment horizon of the asset. Most investors use time-intervals of 10 to 30 years on assets which are perceived as a going concern.

Not all government bonds can be considered risk free and, in some countries, one may have to adjust to find a proper risk-free rate.

### 3.1.8 Market Risk premium

The market risk premium is the level of return between the market portfolio and the risk-free rate. Explained differently, it is the return that investors require to bear the risk that the market portfolio endures. The market portfolio can be both global and country specific, and investors will demand a higher return from country specific market portfolios that exhibits higher risk.

Estimating the market risk premium requires one to predict the future spread between the return of the market portfolio and the risk-free return. While historical data may be used as a future predicator, changing market conditions may render much of the historical data irrelevant. Surveys on a vast number of market-participants are often used to find an estimate for the future market risk premium.

### 3.1.9 Beta

Beta is a measure of systematic risk on a security in comparison with either the entire market or a benchmark portfolio. The beta can be calculated using historical returns so that:

**(9)**

$$\beta = \frac{cov(R_i, R_m)}{Var(R_m)}$$

Where $\beta$ is the beta of the security, $R_i$ is the return of the security and $R_m$ is the return of the entire market or a benchmark portfolio.

As mentioned earlier, investors only care about risk that cannot be diversified away as this would be the most optimal choice in the risk-return world. The benchmark portfolio is a proxy portfolio for all the securities in the entire market. Thus, the benchmark portfolio is a portfolio which have achieved the best level of diversification and should consist of systematic risk only. Assessing the relation between this benchmark portfolio and an individual stock in terms of beta would thus tell us how risky a security is in terms of systematic risk.

When $\beta = 1$, the security inherits the same level of risk as the optimal market portfolio. When $\beta < 1$, the security has less systematic risk than the optimal market portfolio, and when $\beta > 1$, the security has more systematic risk than the optimal market portfolio.

### 3.1.10 Capital Asset Pricing Model

The Capital Asset Pricing Model (CAPM) explains the return of a security in terms of systematic risk and the risk-free rate. The CAPM implies that the following relation for the return of a security is:

**(10)** $$R_i = R_f + \beta_{\mathrm{i}}(R_m - R_f)$$

Where $R_{\mathrm{i}}$ is the return of the security, $R_f$ is the risk-free rate of interest, $\beta_{\mathrm{i}}$ is the return of the security and $(R_m - R_f)$ is the market risk premium.

Three main assumptions underlie the CAPM.

<u>Assumption 1:</u>
Investors can buy and sell all securities at competitive market prices (without incurring taxes or transaction costs) and can borrow and lend at the risk-free rate.

<u>Assumption 2:</u>

Investors hold only efficient portfolios of traded securities; portfolios that yield the maximum expected return for a given level of volatility.

<u>Assumption 3:</u>
Investors have homogenous expectations regarding the volatilities, correlations and expected return of securities.

Assumption 1 may not be realistic for all securities and one can expect to find inefficiencies in the market for some of the securities. Assumption 2 aligns the CAPM relationship with Markowitz (1959)'s capital allocation line regarding the trade-off between return and risk. Assumption 3 may not, when analysing the market, require investors to have the same expectations.

The next subsections cover the single-index model, the multi-factor model, the Fama-French 3 factor model and the Carhart 4 factor model. These are different methods and adjustments to the CAPM which can be done empirically. These methods are made to address flaws with the original CAPM model that comes with the simplicity of the model and bypass some of its rough assumptions.

### 3.1.11 Single Index Model

CAPM is a theory which cannot be directly observed empirically. An approach to estimate the regression relationship on a security without restricting one to the assumptions of CAPM is to use the arbitrage pricing theory by Ross (1976), with the single-index pricing model. The single-index pricing model can be illustrated:

**(11)** $$R_i - R_f = a + \beta_{\mathrm{i}}(R_m - R_f) + \mathcal{E}_{\mathrm{i}}$$

Where $R_i - R_f$ is the excess return of the security, $a$ is the alpha, $\beta_{\mathrm{i}}$ is the beta, $(R_m - R_f)$ is the market risk premium and $\mathcal{E}i$ is the error term.

One can observe the single-index model to not be restricted to CAPM as the single-index model leaves room for alpha-adjusted returns. Alpha-adjusted returns are returns which outperforms the market or benchmark portfolio. By using the single-index model you make the trade-off that the relationship between securities and the market portfolio are not assured.

### 3.1.12 Multi-factor model

To accommodate for further sources of risk, the multi-factor model provides solutions to add more factors that may correlate with the underlying security. Different macro factors that correlates with the security may move in different directions and different magnitude than the total market portfolio. The multi-factor model can be illustrated so that:

**(12)** $$R_i = E(R_i) + B_i F_i + \mathcal{E}_{\mathrm{i}}$$

Where $R_i$ is the return of the security, $E(R_i)$ is the expected return of the security, $B_i$ is the individual factor beta, $F$ is the individual factor and $\mathcal{E}i$ is the error term.

The difficulty with using a multi-factor model is finding relevant factors to use. It is important to find factors that does not suffer from correlation between one another. The multifactor model could be appropriate to use when focusing on a single industry with a common factor, EG: oil price in oil-industry. Utilizing a data sample which consists of several industry would require more variables which could possibly damage the results due to correlation between the factors. Multi-factor model is a collective term, as there are many kinds of those models. One of the most known is the Fama-French three-factor model.

### 3.1.13 Fama-French 3 Factor

The Fama-French 3 factor by Fama & French (1993) is a version of the APT multi-factor model which utilizes two factors, HML (High-minus-low) and SMB (Small-minus-big) when estimating the price of a security. They propose the following relationship:

**(13)** $$R_i = R_f + B_{iM}R_M + B_{iSMB}SMB + B_{iHML}HML + \mathcal{E}_i$$

Where $B_i$ is the individual factor coefficient, $SMB$ is the small minus big factor and $HML$ is the high minus low factor.

The Fama-French 3 factor is in fact just a suggestion of two variables that can be used in the multi-factor model. These two factors have proven to be well used as proxies for external factor and has shown to be uncorrelated between one another, and hence solves one of the main issues of correlation in the multi-factor model.

### 3.1.14 Carhart 4 Factor

The Carhart-Fama-French 4 factor model by Carhart (1997) adds a term to the 3-factor model; the MOM factor, momentum. It captures an anomaly where stocks that have recently had a surge in stock price continues to increase because of the popularity and beliefs of further rise.

It can be expressed as such:

**(14)** $$R_i = R_f + B_{iM}R_M + B_{iSMB}SMB + B_{iHML}HML + \beta_i MOM + \mathcal{E}_i$$

Where $MOM$ is the momentum factor.

## 3.2 Measuring the ESG effects

The measuring of the ESG effect is important, because the measuring will affect the quality and correctness of the analysis. In a perfect world, we would be able to observe the ESG effects perfectly. The world is not perfect however, and many shadow figures exist on the effects of ESG. Doyle (2018) find several biases regarding the reporting practice of ESG effects. For the SRI investor, there are much asymmetric information that might reduce the attractiveness of SRI investments.

Several studies have found that ESG also have geographical differences. Focusing on a single geographic region could account for regional differences which could interfere with the data. Europe has a stricter ESG policy and require companies to be more transparent. Thus, we will only assess data from European companies.

A relatively large sample must also be collected to derive a statistically meaningful analysis. Doing investigative work, minimizing asymmetric risk and finding comparable measures may be too time-demanding and expensive to conduct on its own. A simpler method is to use 3rd party rating agencies that rank companies based on an ESG score. By using 3rd party rating agencies, it assumes that these agencies are providing correct information regarding ESG score. 3rd party agencies, however, may suffer from the same asymmetric information as other investors do.

Regarding the asymmetric information, that may not be an issue. The research question is structured in a way that if investors have access to the same information, asymmetric information will have an insignificant effect[3].

### 3.3 Concept

We start by revisiting the utility function which determines were conventional investors invest on the capital allocation line:

**(15)** $$U = E(r) - 0{,}5A\sigma^2$$

This formula assumes that the only two factors that investors care about is the investments expected return and the risk it ensues. This is referred to as a "risk-

[3] This only holds if hidden information regarding the ESG effects of the firm does not influence stock prices during the estimated period (inside-trade).

reward world." Using the risk-reward world as a base case, investors that are SRI-oriented also care about the impact of their investments. Thus, they derive utility from an additional factor, which is the effect on the public good to the society.

An additional term to the utility function can be added:

**(16)** $$U = U(r, \sigma, ESG)$$

Where ESG is a rating which will be defined in more detail at a later chapter. $\theta$ can be defined as the investor's appetite for ESG effect.

If $\theta > 0$ the investor can be considered a "SRI investor" who derives utility from ESG-friendly investments. As $\theta$ becomes larger, the investor cares more about the ESG effects of the investment and less about the financial returns and risk. We place investors with these preferences in a "Risk-Return-ESG world."

If $\theta = 0$ the investor can be considered a conventional investor which only cares about the effects of return and risk for their investors. Their utility curve can also be defined as the normal utility curve. We place investors with these preferences in a "Risk-Return world".

If $\theta < 0$ the investor can be considered as a "Sin investor" who prefers to invest in stocks which apply damage to the society. Such behaviour may be considered mentally deranged and investors classified in this category are few and far between in frequency and small in significance (in relation to invested capital). Considering this, we assume that $\theta$ will not have a negative value.

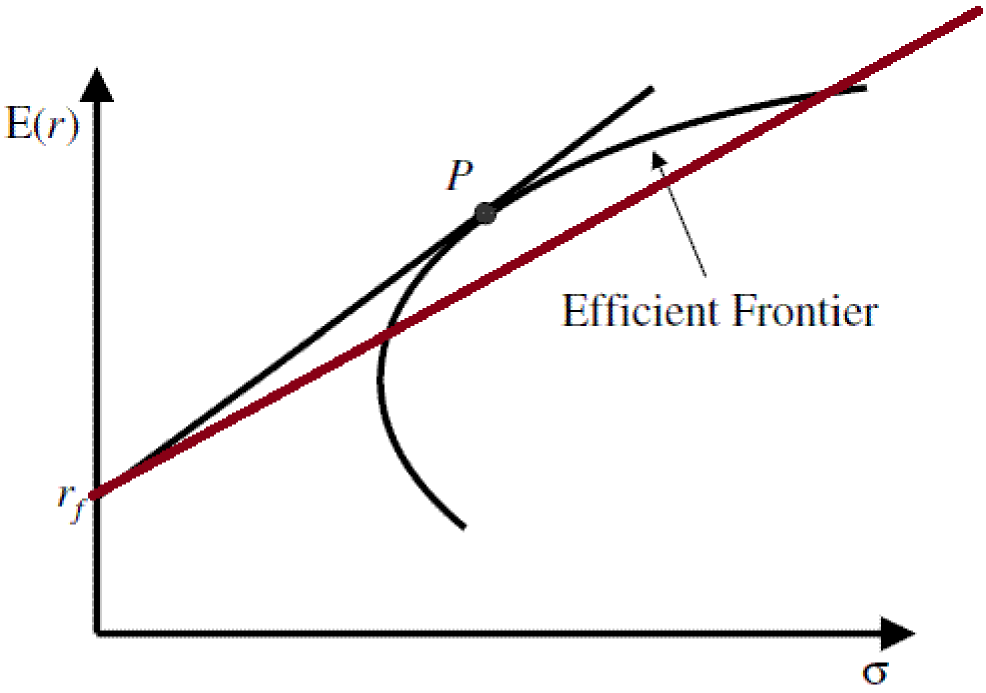
We also recognize a problem with the utility curve when $E(r) < 0$, $\sigma^2 = 0$ and yet the utility combination is the most optimal. A combination like this is called a donation and is not meant to collect a financial return.  Donations may present a problem in our assumptions towards the linearity of the preference of ESG. If investors perceive to derive more utility per "unit of currency" when donating rather than investing, a special case of mental accounting bias may influence the analysis. (Ackert & Deaves 2009) The bias is not expressed in monetary terms here however, but in utility towards ESG related investments. Another explanation could be that investors do not trust the impact that the SRI investments provide. If that is the case, an asymmetrical information cost is applied to the SRI investments.

In accordance with the property of the proposed utility curve, we believe SRI investors in a "risk-reward world" will overinvest in stocks with ESG values. Another problem SRI investor may face, however, is the effect of diversification on the portfolio. If their total portfolio excludes a significant amount of stocks (due to the negative screening of sin stocks while also placing a significant amount in positive screened SRI stocks), their portfolio may be affected by both systematic and unsystematic risk.

Revisiting Markowitz's (1959) capital allocation line, investors in a "risk-reward" world will want a portfolio which is tangent with the capital allocation line where the optimal return/risk combination exist. This portfolio is properly diversified and does not suffer from unsystematic risk. The combination, however, may not be the optimal portfolio for an investor in the "risk-reward-ESG world". If the optimal portfolio in a "risk-reward world" is also not the combination that provides the most value to SRI investors, the two portfolio combinations will be different.

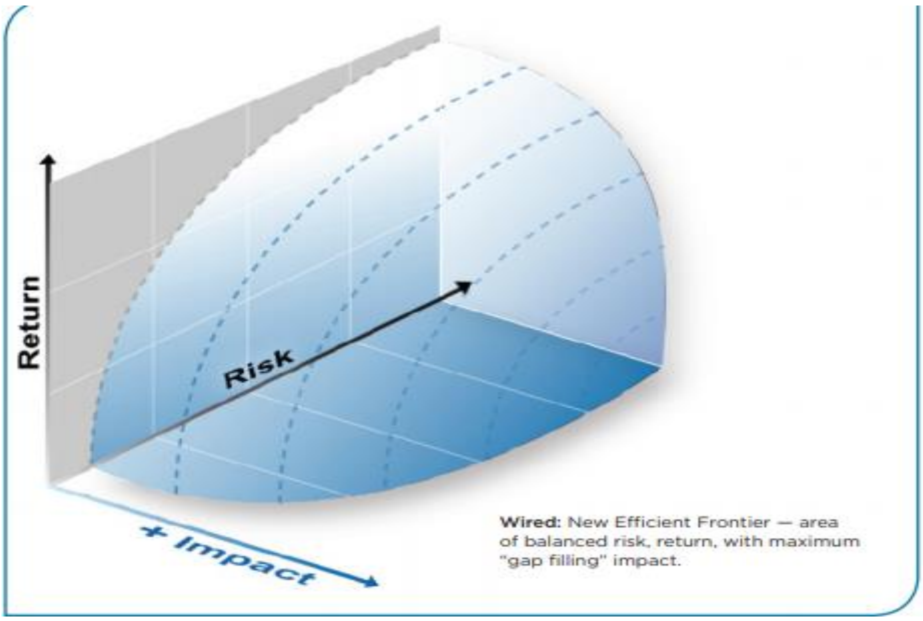
Illustrated in the traditional capital allocation line from section 3.1.5:

*Figure 2: Non-optimal Capital Allocation Line in a risk-reward-ESG world.*




But this graph does not capture the adequate relationship that we suspect the SRI investor to inherit. A more proper illustration would be to illustrate a newly proposed capital allocation line as such:

*Figure 3: The "new" Capital Allocation Line*




Source: Impactassets.org

In this figure, we capture not only the relationship of the conventional investor but also the relationship for investors with SRI preferences. We ignore negative values of $\theta$ as we assumed earlier that no investors will have such values.

It is common practice to use funds when estimating regression analysis as the returns are more consistent and lenient to shocks. As discussed in the research section, "SRI" funds have been prone to biases and tends to slide towards the same characteristics and qualities of conventional funds. Further, it could be problematic to analyse an SRI fund in terms of standard fund performance measures (Eg: Alpha, Treynor) if the SRI fund is prone to unsystematic risk. This is because using such measures assume the funds to have no unsystematic risk. Individual stocks will thus be used for analysis. To handle the issue of consistency, a rather large sample will have to be collected to smooth the regression analysis.

Recall the capital asset pricing model (CAPM):

**(10)**
$$R_i = R_f + \beta_\mathrm{i}(R_m - R_f)$$

The capital asset pricing model is derived from the capital allocation line. The CAPM assumes that investors will only have the traditional utility curve with preferences for return and risk, thus, investors with a preference for ESG should theoretically underperform in relation to its conventional counterpart.

The single-index model only holds for the conventional investor in a "risk-return" world. A proposed model for the single-index model in the "risk-return-ESG" world would be to add another term to the equation such that:

**(17)**
$$R_i = R_f + \beta_\mathrm{i}(R_m - R_f) + \theta ESG$$

The two terms can, due to the CAPM assumption on unrestricted short-sales, not be compared. The CAPM model with the additional ESG term assumes that all investors have the same preferences for $\theta$, the appetite for ESG. The traditional CAPM model, however, assumes that individual investors can have different preferences of risk, yet all investors will want to hold the same portfolio of stocks.

Considering that, theoretically, SRI investors will hold different portfolios which performs sub-optimal in a traditional CAPM world to the common portfolio that is held by conventional investors. As $\theta$, the appetite for ESG, can be considered a linear parameter, more preference for ESG should translate to lower returns.

Is the single-index model with the ESG extension appropriate to use for explaining the ESG factor? The Fama-French-Carhart four factor model proposes factors which covariate with the small-company bias, high book values and a momentum bias. One is not limited to only using the single-index model or the Fama-French-Carhart 4 factor model for analysis and it is possible to compare and discuss the results of these two models.

The Fama-French-Carhart four factor model with ESG estimation would receive an extension so that:

**(18)** $R_i = R_f + \beta_{iM} R_M + \beta_{iSMB} SMB + \beta_{iHML} HML + \beta_i MOM + \theta ESG + \mathcal{E}_i$

A quick glance on the model would show that this model is like the single-index model, except adjusting for some marked biases. If the factors do not correlate with the ESG factor, the ESG coefficient should stay the same for both models.

3.4 Model specification and estimation

When Gauss-Markov assumptions are unviolated, the ordinary least squares (OLS) regression is the most efficient estimator. The OLS regression minimizes the sum of the squared residuals of a model.

**(19)** $$y_{it} = \beta_0 + \beta_1 + \beta_2 x_{it} + .. + \beta_3 x_{\mathrm{i}}(t + k) + \mathcal{E}_{it}$$

Where $y_{it}$ is the dependent variable, $\beta_0$ is the constant, β(1, 2, k) is the independent variables and $\mathcal{E}_{it}$ is the error term. The Gauss-Markov assumptions are a set of assumptions which requires linearity, random selected sample, non-collinearity, exogeneity and homoscedasticity. Our data includes several companies and are sampled over time, which means that our data is panel data. The panel data version of OLS is pooled-OLS and is based on the same method as the regular OLS. Pooled OLS ignores differences in time so that all observations are treated as they are collected in the same time, t. In regular OLS, one may assume that the sample are independent from each other. In pooled OLS however, correlation between the error terms may occur as observations are sampled multiple times from the same individual. The correlation would lead to a bias in the estimator which would mean that it would no longer be the best linear unbiased estimator (BLUE).

The error term, v (also called the composite error) can be divided into two factors, $v_{it} = a_i + u_{it}$ . $a_i$ capture the unobserved time-constant effect that affects $y_{it}$, while $u_{it}$ is the idiosyncratic error, the error which varies over time and affects the error term. Even though $u_{it}$ (which is also the composite error term in cross-sectional time series) is uncorrelated with the independent variables, the time error $a_i$ may affect the independent variable and thus cause a biased called the heterogeneity bias. To account for this unobserved time-constant effect, it is possible to do an alternate approach called the fixed effects estimation.

Assume that one has the equation:

**(20)** $$y_{it} = \beta_1 X_{it} + a_i + u_{it}\,, \qquad t = 1, 2, \dots, T,$$

Then, calculate the time-average of the equation:

**(21)** $\bar{y} = \beta_1 \bar{X}_{\mathrm{i}} + a_i + \bar{u}_{it}, \quad t = 1, 2, ., ., T,$

Then by subtracting Equation **(21)** from Equation **(20)** you get:

**(22)** $\ddot{y}_{it} = \beta_1 X_{it} + u_{it}, \quad t = 1, 2, \dots, T$

Observing equation **(22)**, the time-constant variable a has been removed and only the idiosyncratic part of the error term, which we assume to be uncorrelated with the independent variables, remains. Under this assumption, the estimator is unbiased. The equation over assumes a simple-regression model, but there are no changes in the underlying procedure by adding more variables to the equation. The problem by doing a fixed-effects procedure is that one cannot have other time-constant variable as the variable would be removed when subtracting equation **(21)** with equation **(20)**.

Random effects are another estimation method which can be used to bypass the effects of time-constant error term.

The starting point:

**(23)** $y_{it} = \beta_0 + \beta_1 X_{it} + .. + \beta_k X_{itk} + a_i + u_{it}, \quad t = 1, 2, \dots., T,$

We assume here that the unobserved effect a has a zero mean (we would not be able to include the intercept if not). We further assume that the fixed effect now is uncorrelated with each of the independent variables:

**(24)** $Cov(X_{itj}, a_i) = 0, \quad t = 1, 2, .., T; j = 1, 2, .., k.$

Running a pooled OLS under these conditions could still make it biased. When doing this estimation, you still risk correlation over time between the idiosyncratic error and the time-constant error:

**(25)** $Corr(V_{it}, V_{is}) = \frac{\sigma^2 a}{(\sigma^2 a + \sigma^2 u)}, \; t \neq s$

In cases where serial correlation is significantly large, this would have an impact on the estimation results. To account for this, the same logic as the fixed effects is applied, but instead of subtracting the equation by its “time”-mean, you construct a term for the correlation so that **(25)** becomes:

**(26)** $$\theta = 1 - \left[\frac{\sigma^2 u}{(\sigma^2 u + T\sigma^2 a)}\right]^{1/2}$$

Equation **(26)** is then subtracted with **(25)** and becomes:

**(27)** $$y_{it} - \theta\bar{y}_i = \beta_0(1-\theta) + \beta_1(X_{it1} - \theta\bar{X}_{i1}) + .. + \beta_k(X_{itk} - \theta\bar{X}_{ik}) + (v_{it} - \theta\bar{V}_i)$$

While the fixed effects estimator subtracts the mean, the random effects only removes a fraction of the time average mean. This transformation accounts for the serial correlation and allows for other time-constant variables to be estimated.

Now which estimator would be appropriate? When Gauss-Markov assumptions are satisfied, regular OLS, or in this case pooled OLS, would be considered BLUE. To check if pooled OLS is appropriate, one can use the Lagrange multiplier-score test. The score test estimates if there could exist an endogeneity problem when pooling the sampled panel-data. Since random effects model partially negate the time effects while fixed effects remove the entire mean, random effects are more efficient. Random effects however assume the error term to be uncorrelated with the independent variables. If this assumption does not hold, this estimation method is not appropriate to use. The Durbin-Wu-Hausman test can be used to test for differences of the Fixed effects and Random effects model. It tests for differences between a consistent estimator and an inconsistent estimator by its variance to see if the inconsistent estimator is appropriate to use. If there are no significant differences, random effects model is also considered a consistent estimator.

Some methods and difficulties in selecting our estimator has now been covered. We will now discuss our data sample and other factors which will affect the choice of estimator. In terms of efficiency, assuming unbiasedness, OLS > RE > FE (where OLS is the most efficient).

The data collected is panel data which means that we will have to use the Lagrange-multiplier score test to check if running a pooled OLS is appropriate. If results from the score test show that it is appropriate, a pooled OLS will be used because it would then be BLUE.
If score test shows that it is not appropriate, we will discuss the random effects estimator. The random effects estimator requires the estimation of the fixed effects estimator to compare the volatility. An underlying problem with doing so is that it restricts the use of independent variables with constant values over time values, t. Not only does that permit the use of dummy variables which can find differences in country, industry etc, but the ESG scores for the firm are held constant over the entire time period. This assumption in its own is not true, but the collection of ESG score are a new phenomenon and only a handful of companies have data which stretches back for 10 full years. Reporting ESG scores are not enforced by law and thus many companies also have not fully updated ESG scores for each year. Luckily, ESG scores does not vary much over time and thus, an assumption of constant ESG score over the estimated time period may not have major effects on the underlying data sample. Only including companies which fulfils the requirement of 10 full years of reporting practice would weaken the sample considerably.

The constant ESG scores are also a reason why the first-differences approach has not been mentioned as an estimation technique. It would not be possible to use this approach when there is a time-constant time variable. Fixed-effect estimator is mentioned because it plays a part in estimating the Random-effect estimator.

If having a constant ESG term, why not use cross sectional data? The inherent properties of our regression restrict us from using cross sectional data. To explain further, the benchmark index would have the same average return for all observations, thus violating Gauss-Markov assumption and causing the OLS regression to not be possible anymore.

An imperfect solution is to not include the ESG term nor the dummy variables when estimating the fixed effect estimator. This means that the volatility of the dummy variables and the ESG variable may be inconsistent. It however asserts the validity of the rest of the model.

We ran a Breusch and Pagan Lagrangian multiplier test for random effects to check if pooled OLS was appropriate to use. The test gave statistical differences in data which shows that pooled OLS was not appropriate. [4]When comparing the fixed-effects model with the random effects model (not including the ESG term), we find no significant differences and thus the random effects model without the ESG term and coherent dummy variables is efficient and consistent to use.

Is random effects model with the ESG term and the dummy variables appropriate to use then? Technically, we cannot know if they are consistent when estimating the model. Doing the comparison of the fixed effects model and the random effects model without the terms only limits the potential pitfall but it does not eliminate it. Furthermore, constant ESG scores is risky to use if the scores differ significantly over time. When reading into the results of this analysis, the reader should be aware of the potential pitfalls that the results may have.

For our estimator, the random-effects estimator will be used. Our reasoning is based on moments from the text above and from the outcome of the tests. We acknowledge the risks of using this method, but we also desire the qualities of the estimator which helps us bypass the bias a pooled OLS would have.

## 3.5 Analysis

The CAPM formula and our proposed CAPM with an extra ESG term is not an empirical formula. By adding another term to the single-index model from equation **(11)** the new equation is:

[4] See Appendix 2

$$(28) \qquad R_i - R_f = a + \beta_i(R_m - R_f) + \theta ESG + \mathcal{E}_i$$

Furthermore, the Fama-French-Carhart’s four factor model empirical version with an extra ESG term would look like this:

$$(29) \quad R_i - R_f = a + \beta_i(R_m - R_f) + B_{iSMB}SMB + B_{iHML}HML + \beta_i MOM + \theta ESG + \mathcal{E}_i$$

The formulas will be regressed as a random-effects model in line with our results from the previous section.

## 4.0 Data

This chapters gives a brief insight into the various data-components of our regression. Most of the data have been collected using the Thomson Reuter Datastream software. By using the charting function and the ESG-scores function, we could accurately and efficiently extract the data we needed.

We decided to collect data from a 10-year period. The reason we use this estimation period is to get accurate estimates of the individual companies’ performance over time. The estimation period used in this analysis is 31.12.2008 to 31.12.2018. Furthermore, we collected monthly data, except for the ESG Scores. Yearly data would in our opinion be too infrequent, and daily or weekly data could potentially become too noisy with large swings.

### 4.1 Individual stocks

The monthly returns were gathered from a sample of 928 individual companies across 16 countries in Europe: Austria, Belgium, Denmark, England, Finland, France, Germany, Ireland, Italia, the Netherlands, Norway, Poland, Portugal, Spain, Sweden and Switzerland. The sample is collected from country-specific lists over companies that have available information on their ESG scores. All the countries are also included in the STOXX Europe 600 Index.

Several companies from the country-specific lists did not have data for the entire estimation period, and hence have been identified and omitted from the sample.

### 4.2 ESG Scores

ESG scores are collected from the Thomson Reuters ESG Scores. Over 150 content researchers provide information for ESG ratings of companies around the globe. There are ratings available for over 7 000 companies. The ratings for ESG scores are divided into environmental, social and governance factors which factors are further divided into 10 categories which they are ranked at.

ESG scores are rated after a set of score-measures which have its strengths and weaknesses. The consistency of the ESG score-collecting method is a strength which would improve the comparability of the scores but could also lead to companies "chasing" score benchmarks rather than just behaving in a socially responsible way. This could further mean that some companies are higher rated than they deserve.

### 4.3 Benchmark portfolio

The STOXX Europe 600 was chosen as the benchmark portfolio. STOXX 600 consists of around 90 % of the market capitalization of Europe. (Stoxx.com) As we have only collected data from selected countries in Europe, and all these countries are included in the STOXX 600, it seems fitting to use as benchmark given its cover of market capitalization and the European region.

### 4.4 Risk Free rate

The risk-free rate was gathered from the European Central Banks measure for government AAA spot rate 10-year bonds. Returns were collected for the estimation period in monthly intervals.

### 4.5 Industry

Thomson Reuter has an additional function called Thomson Reuter Business Classification (TRBC). TRBC classifies global companies by industry. The classification consists of four levels of hierarchical structure from top to bottom: Economic sector, business sector, industry group and industry. Each individual company is listed with their specific industry, and we used a business classification table published by Thomson Reuter to trace each industry back to its coherent economic sector. All the companies were placed within one of ten different economic sectors: Energy, Basic materials, Industrials, Cyclical consumer goods & services, Non-cyclical consumer goods & services, Financials, Healthcare, Technology, Telecommunication services and Utilities.

## 5.0 Results and discussion

In this chapter we run all the relevant regression and present our results. We include results from 10, 5, 2- and 1-year data. Additionally, we include a regression including sector and country-dummies. We interpret and discuss the results and try to draw inferences by looking at each regression, and also by comparing them. The base-case regression is run by regressing Equation **(28)** and **(29)**.

### 5.1 "Base case"

*Table 1: Regression results with 10-year data.*

| | Single-Index Model | FF - Carhart 4 Factor |
|---|---|---|
| Risk premium | 1.0100*** (0.0102) | 1.0103*** (0.0121) |
| ESG | -0.1228*** (0.0453) | -0.1228*** (0.0453) |
| HML | | 0.1978*** (0.0197) |
| SMB | | 0.6189*** (0.0242) |
| MOM | | -0.1236*** (0.0121) |
| Constant | 0.0131*** (0.0029) | 0.0126*** (0.0029) |
| R^2 | 0.0801 | 0.0892 |
| Wald Chi | 9848.36 | 11072.56 |

| | | |
|---|---|---|
| P-value | 0.0000 | 0.0000 |
| Observations | 111 359 | |
| Groups | 928 | |

*ESG scores are presented in thousands. *** indicates significance at a 1% level.*

The single index model predicts the constant to be 1.31%, the beta of the market to be approximately 1.01 and each increase in ESG score to reduce the expected monthly return of the company with 0.01228 %. The model is significant at a 1 % level, both for individual variables and for the model wholly. The explanation power is low, which tells us that the variables only covers a fraction of the estimated returns in aggregate.

The Fama-French-Carhart Four factor predicts the constant to be lower, at 1.26 %. The beta of the market is predicted to be approximately 1.01, very close to the prediction of the single-index model. A 1 % increase in the SMB or HML would predict an increase in returns by 0.61 % and 0.19 % respectively. A 1 % increase in the MOM factor would predict a decrease in returns by 0.12 %. The ESG factor is estimated to be the same as the single-index model, decreasing predicted returns by 0.01228 %. This model is also significant at a 1 % level for the individual variables and the model. The explanation power has increased, but only slightly. This tells us that the additional factors introduced in this model do not contribute much to explaining the dependent variable return.

According to the logic of single-index model and the Fama-French-Carhart Four factor model, theoretically by holding this data sample as a portfolio, investors should be able to achieve *yearly* constant-returns of 15,72 % and 15,08 % when annualizing (multiplying with 12) the constant in the model above. The interpretation of the model changes when adding the ESG variable to the normal single-index model and the Fama-French-Carhart Four factor model. Interpreting and comparing the constant as alpha-return of an original model would be a mistake. The interpretation of alpha-return is the excess return of an investment in relative to the return of a benchmark index, not in relative to both the benchmark index and an ESG factor. Furthermore, assuming 15.72 % or 15.08 % return of an investment when beta = 0 would also assume the ESG score to be 0, which is over its lower limit (ESG scores rank from 1-100). This massive constant is also unclear as the lowest reported ESG score of a security is 13. The effects ESG score has on returns with ratings lower than 13 is thus unclear.

The coefficient to the benchmark portfolio for both models are very close to 1. One could assume that the data is thus properly sampled so that the underlying data reflects its coherent benchmark. Our research question is to observe if SRI-oriented investors care about the level of "ethicalness" of a security. The reported ESG variable for both single-index model and the Fama-French-Carhart Four factor reveals that for each increased point of rating in terms of ESG will reduce the predicted monthly return of a company with 0.01228 %, or *Yearly* 0.14736. Table 2 presents how the underlying ESG scores are distributed:

*Table 2: Underlying ESG-scores distribution.*

| | ESG |
|---|---|
| Mean | 61 |
| Lowest | 13 |
| 10 percentile | 39 |
| 25 percentile | 51 |
| Median | 63 |
| 75 percentile | 72 |
| 90 percentile | 79 |
| Highest | 96 |

To assess the impact of the ESG variable, the ESG scores are limited to a maximum of 100, which is the top score. It can also be observed from percentiles that the data is unevenly distributed. Moving from 25 percentile to 75 percentiles would ceteris paribus predict a lower *yearly* return of 3.10 %. The effect of SRI gets even more severe as you compare investments that fall in the lower 10 percentile with the 90 percentiles. The predicted returns then fall with 5.89 %. Investing based on ESG scores are not only statistically significant, but also economically significant.

Would these results answer our topic question? Not necessarily.

## 5.2. Time-intervals

One major assumption made earlier was ESG scores to be assumed constant over the entire period. For an estimation period of 10 years, these results can be distorted by the fallacy of such assumption. A way to deal with this is to exclude earlier years, effectively limiting the estimating period. By doing so, samples will be shortened but their relevance will increase. Estimation intervals will be set at 5, 2 and 1 year prior to see if there are changes to the reported variables.

*Table 3: Regression results with 5-year data (2014-2018).*

| | Single-Index Model | FF - Carhart 4 Factor |
|---|---|---|
| Risk premium | 0.8367***<br>(0.0116) | 0.9224***<br>(0.0125) |
| ESG | -0.0000<br>(0.0000) | -0.0000<br>(0.0000) |
| HML | | 0.0903***<br>(0.0225) |
| SMB | | 0.6105***<br>(0.0269) |
| MOM | | -0.0383<br>(0.0209) |
| Constant | 0. 0101***<br>(0.0018) | 0.0096***<br>(0.0018) |
| R^2 | 0.0859 | 0.0948 |
| Wald Chi | 5253.58 | 5855.36 |
| P-value | 0.0000 | 0.0000 |
| Observations | 55 740 | |
| Groups | 928 | |

**** indicates significance at a 1% level.*

When changing from 10-year estimation period to 5-year results change significantly. ESG scores are no longer statistically significance and thus inferences can no longer be drawn from its results. Results from Single-Index model and FF-Carhart 4 factor are also more different, where the beta of the market risk premium goes from 0.84 to 0.922 when using the 4-factor model. In comparison, when using the entire data sample, the change between the two models of the estimated beta is economically insignificant.

*Table 4: Regression results with 2-year data (2017-2018).*

| | Single-Index Model | FF - Carhart 4 Factor |
|---|---|---|
| Risk premium | 1.0062***<br>(0.0246) | 1.0326***<br>(0.0249) |
| ESG | -0.0000<br>(0.0000) | -0.0000<br>(0.0000) |
| HML | | 0.0001<br>(0.0414) |
| SMB | | 0.6007***<br>(0.0623) |
| MOM | | 0.0460<br>(0.0494) |
| Constant | 0.010***<br>(0.0106) | 0.0106***<br>(0.0031) |
| R^2 | 0.0691 | 0.0744 |
| Wald Chi | 1678.66 | 1816.59 |
| P-value | 0.0000 | 0.0000 |
| Observations | 22 272 | |
| Groups | 928 | |

**** indicates significance at a 1% level.*

Results when using only 2 year of the sample shows the ESG variable to exhibit more significance than when using 5 years prior. The results from the two models vary, which means that there exists some correlation between some of the Carhart-Fama French variables and the ESG variable. The beta for the single-index model and the Carhart-FF 4 factor is 1,01 and 1,03 respectively, which means that the average of the analysed stocks represents more closer the systematic risk of the benchmark-index.

*Table 5: Regression results with 1-year data (2018).*

| | Single-Index Model | FF - Carhart 4 Factor |
|---|---|---|
| Risk premium | 1.0010***<br>(0.0115454) | 1.0437***<br>(0.0317) |
| ESG | 2.96e-06<br>(0.0000) | 2.96e-06<br>(0.0000) |
| HML | | -0.0662 |

| | | |
|---|---|---|
| | | (0.0681) |
| SMB | | 0.5889*** (0.1408) |
| MOM | | -0.0149 (0.0747) |
| Constant | 0.0030 (0.0041) | 0.0060 (0.0043) |
| R^2 | 0.0835 | 0.0886 |
| Wald Chi | 1030.45 | 1100.28 |
| P-value | 0.0000 | 0.0000 |
| Observations | 11 136 | |
| Groups | 928 | |

**** indicates significance at a 1% level.*

The 1-year analysis shows consistency in the beta variable, but the ESG variable is very insignificant and shows ESG to have no effects on the profitability of the company. From our original base case we have gone from having both statistical and economical significance to having either. The analysis of excluding samples has not only failed to uncover any problems on our assumption of constant ESG scores, it has also revealed an even more dangerous issue to our results.

It would be easier to conclude ESG variable to have no effect if there was consistency in these results. When analysing the first 5 years (from 2009 up to 2014) with data, we receive significant results for ESG variables again. The results can be shown in appendix 1. One could think that the significance of the ESG variables are based upon luck of the draw, that is, the results depends randomly on the time period you have chosen to collect the data. We do not believe the varying significance of the ESG variable to be purely random.

When boiling down to microeconomic terms, the effects of ESG could be a combination of both a public and a luxury good. A luxury good would mean that in times of high income, ESG effects are more preferred than in times with low income. In the aggregate picture, investors could thus be assumed to command high income when enduring times of economic boom. If this is the case, further analysis should control for economic boom and acknowledge the investors utility function of ESG to not be linear but to depend on several factors.

Nevertheless, as time progresses and more data is collected for ESG levels, the business cycle will become less noisy and will be easier to gather results from. Given our limited data sample of 10 years back (and we even assume ESG scores to be constant to get more data!), having only one or two changes in the cycle as historical data may not be enough to control for its effect and thus will probably not be able to linearize the ESG variable.

The differences in results over the estimated time period are worrisome and damages the legitimacy of the results on our 10-year data collection. We can however still draw consistent conclusions towards the ESG score. It may be better to use the 10-year period because it will better capture times of expansions/booms and times of contractions/recessions.

### 5.3 Accounting for nation and industry differences

When assessing the effects of ESG we can also account for both nation and industry differences to see if there exist significant differences which can influence the interpretation of our results. We constructed dummy variables for 16 countries and for 10 industries. To test the significance of the dummy variables we employed an elimination method where we removed the least significant dummy variable and then repeated such procedures until we ended up with only significant dummy variable. This procedure is to remove the covariation effect that an insignificant variable possesses over a significant one.

We employed the elimination technique, where we remove the least significant variable for the countries first and found France and Italy to have a significant effect on the regression analysis. Only the energy sector showed to have a significant effect when employing the elimination technique on industry variables. When combining the original Carhart-FF 4 factor with the three significant dummy variables we receive results which are presented in the table below:

*Table 6: Regression results with industry and country dummies included. Only the significant dummies using the elimination technique are included due to lack of space.*

| | FF - Carhart 4 Factor with significant dummy variables | FF – Carthart 4 Factor with all dummy variables |
|---|---|---|
| Risk premium | 0.8993*** (0.0122) | 1.0104*** (0.0122) |
| ESG | -0.1375*** (0.0000) | -0.1385*** (0.0001) |
| HML | 0.1996*** (0.0198) | 0.1978*** (0.0197) |
| SMB | 0.5852*** (0.0243) | 0.6189*** (0.0242) |
| MOM | -0.2041*** (0.0122) | -0.1236*** (0.0121) |
| France | 0.0055** (0.0024) | 0.0114 (0.0071) |
| Italy | -0.0080 (0.0033) | -0.0013 (0.0075) |
| Energy | -0.0058** (0.0029) | -0.0024 (0.0187) |
| Constant | 0.0300*** (0.0029) | 0.0039 (0.0200) |
| R^2 | 0.0802 | 0.0901 |
| Wald Chi | 9819.88 | 11110.27 |
| P-value | 0.0000 | 0.0000 |
| Observations | 111 359 | 111 359 |
| Groups | 928 | 928 |

*ESG-score is presented in thousands. *** indicates significance at a 1% level.*

Adding the dummy variables changes the magnitude of the ESG variable from - 0.01228 % to - 0.01375 % monthly which means that regional differences and industry have some effect on the impact of ESG variables. The differences in nations could be because of informational inefficiency or some asymmetric reporting practice, but this is speculative.

Furthermore, when adding all variables, significant or not, changes the ESG variable with - 0,0001 %.

The energy industry variable is grouped rather unfortunate because it combines the effect of oil-related industries and renewable energy as both types of companies fall in the same economic sector. SRI oriented investors with a focus on climate would invest in some stocks and refrain from investing in other stocks in the energy dummy variable. While we have assumed ESG scores to be adequate for the purpose of our analysis, there is no denying that clean energy is something that SRI investors will prefer over less clean energy. Thus, grouping unclean and clean energy together as a control variable could lead to ESG scores being negatively biased, leading to an overprediction of the negative effects of high ESG scores.

## 6.0 Conclusion

The research question we presented were if SRI investors consider the ESG effects of their investments. After conducting our base analysis, we end up with results that show ESG securities to underperform relatively to the benchmark index. With these results in mind we hypothesize that the underperformance may stem from overinvestment in said securities because of the utility they provide for morally driven investors.

Earlier we assumed ESG scores to be constant. We explored this assumption by shortening the time period of our collected data sample. If our assumptions towards ESG scores was correct, then the variables would ceteris paribus, remain unchanged. The variables did not only change, but they turned out statistically insignificant. Only the first 5 years (2009-2014) of the data sample turned out to be significant when dividing the data. We do not believe the inconsistency of such results to be from the luck of the draw, rather we believe it to be structural; dependent on the business cycle.

An extra addition to the analysis presents dummy variables which accounts for significant differences in both country and industry sector. After checking for significance, we end up finding significant dummy variables on France and Italy, and the Energy sector.

Our results are consistent with our preliminary predictions. Using a 10-year data set which includes dummy variables for France, Italy and the Energy sector, a 1-point decrease in the ESG score is associated with a decreased annual return of 0,165%. The magnitude of this becomes clear when assessing the values from the 10 % "worst" companies with the 10 % "best" companies in terms of ESG, with a prediction difference of whoppingly 10,89 % annual return. (Cf Table 2)

We must be careful to conclude this as if it is the investors willingness to take a financial loss for improved ESG performance. Not only have results been inconsistent, but they may also be biased if ESG is correlated with investments in clean energy and our energy variable probably fails to control for this effect as it is mixed with "unclean" energy sources. To add even more confusion and inconclusiveness, we cannot safely conclude if ESG investments are done in the name of altruism or whether ESG practice are a product of marketing and public relations.

Compared to many other studies however, we do find economically and statistically significance on the ESG variable. The vagueness of the SRI definition when assessing former studies may have led to different interpretations of the term, and thus may have caused differences in results. Our study gains credibility by assessing well defined ESG scores from a trusted 3$^{rd}$ party. Our model is consistent with former logic and assumptions made are robust for more extensions to the models proposed. Consistency can also be observed in this analysis when comparing the single-index model results with the Fama-French-Carhart model results.

It is important to re-mention that our findings do **not** imply SRI investors to behave sub-optimally or irrational. Our assumptions and predictions imply that SRI investments perform worse financial wise because they are more attractive for investors. In other words, investors on average, derive more utility by investing in stocks which are perceived by the common eye to be more ethical. For the company's management, having a higher ESG rating should translate to a more attractive company which have access to more capital.

## 7.0 Further research

Our thesis opens for further research, many which are alterations and additions of techniques which we have used. The thesis follows procedures which may not provide full clarity to the issue we are researching and does make assumptions which may not be always correct.

Constant ESG scores are assumed for the entire period, with the latest reported ESG score as the constant. Conducting an analysis which lets ESG scores vary over time could change results if the assumption is false. The intention of letting ESG scores vary over time is to let them be more correlated with earlier dates in the sample. A counterintuitive point, at least to this sample, is that there is more significance towards the ESG variable at earlier dates of the sample. This relationship could be spurious, but most likely is the assumption towards constant ESG score not detrimental for the significance of the analysis. Nevertheless, it could be interesting to investigate how the results would change.

When we shortened the time period, problems revolving statistical significance was encountered. A potential explanation of this could be due to change in business cycles. Investors may have a higher preference of high ESG-score investment during times of economic boom rather than during times of recessions or contractions. Further analysis could apply measurements which can track business cycles and thus control for the effect of the preference of high ESG score. Ideally, an analysis containing a business cycle variable should be able to linearize the ESG variable. In other words, the variable should covariate with the ESG variable and control for the effects of changing market conditions on the preference of investments with high ESG score.

SRI investing is wide, while the ESG score rating is based on a set of definition of a rule-based scorecard, SRI investors may thus defer from investments that have a high ESG-score rating if the investment goes against their moral compass. Oil companies in this sample have received high ESG score rating because they follow good social and governance practices, as well as investing in environmental issues. Equinor for example, have a score of 82, despite having most of their business related to oil and gas services. Environmental driven SRI investors may refrain from investing in Equinor because of this, even though it scores high.

Dividing industries down in smaller segment could help separate this effect. For the 10 industries which was chosen, only the energy industry showed up significant. Differences in the energy industry (clean/unclean energy) could have implications on the results. Unfortunately, the sample will get severely limited by doing so and some of the industry groups would only contain a single company. Adding more companies to the sample by for example not limiting one to Europe could help distinguish between the effects of clean and unclean energy.

## 8.0 References


Markowitz, Harry. «*Portfolio selection.* » Research paper, American finance association, 1952. Accessed at: https://www.math.ust.hk/~maykwok/courses/ma362/07F/markowitz_JF.pdf.

Sharpe, William. *"Capital asset prices: A theory of market equilibrium under conditions of risk."* Research paper, American finance association, 1964. Accessed at: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1540-6261.1964.tb02865.x

Jensen, Michael C. *"The performance of mutual funds in the period 1945-1964."* Research paper, Harvard Business School, 1968. Accessed at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=244153

Roll, Richard. *"A critique of the asset pricing theory's tests Part I: On past and potential testability of the theory."* Research paper, University of California, 1977. Accessed at: https://www.sciencedirect.com/science/article/pii/0304405X77900095

Ross, Stephen A. *"The arbitrage theory of capital asset pricing."* Research paper, University of Pennsylvania, 1976. Accessed at: https://www.sciencedirect.com/science/article/pii/0022053176900466

Banz, Rolf W. *"The relationship between return and market value of common stocks."* Research paper, Northwestern University, 1980. Accessed at: http://www.business.unr.edu/faculty/liuc/files/BADM742/Banz_sizeeffect_1980.pdf?fbclid=IwAR2zulAphLo3IyeTuiu1_1dT_GYlnD4m_pLmnHnAFJFmrQmwEDEekmDY8PU

Elton, Edwin J, Gruber, Martin J, Das, Sanjiv & Hlavka, Matthew. *"Efficiency with Costly Information: A Reinterpretation of Evidence from Managed Portfolios"* Research paper, The review of financial studies, 1993. Accessed at: https://academic.oup.com/rfs/article-abstract/6/1/1/1575864?redirectedFrom=fulltext

Lee, Darren D., Humphrey, Jacquelyn E., Benson, Karen L. & Ahn, Jason Y. K. «*Socially responsible investment fund performance: the impact of screening intensity.* » Research paper, UQ Business School, University of Queensland, St Lucia, 2010. Accessed at: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1467-629X.2009.00336.x

Bauer, Rob, Otten, Roger & Rad, Alireza Tourani. *"Ethical investing in Australia: Is there a financial penalty?"* Research paper, Maastricht University & Auckland University of Technology, 2006. Accessed at: https://sustainable-finance.nl/upload/researches/Bauer-et-al_Ethical-Investing-in-Australia.pdf?fbclid=IwAR3lo_iMlddhn9ORIeAcQsNvZrIl_at4J9KTOihsqdp606YgpNsa-vXZgbw

Schröder, Michael. *"The performance of socially responsible investments: Investment funds and indices."* Research paper, Centre for European Economic Research, 2004. Accessed at:

https://www.researchgate.net/publication/227323290_The_Performance_of_Socially_Responsible_Investments_Investment_Funds_and_Indices?fbclid=IwAR1NAMfOvLNn3g_LKTNohN-3Ix0PTdF4Xqi0On5TNXiSB_AFmY8NVklTgAc

Bello, Zakri Y. *"Socially responsible investing and portfolio diversification."* Research paper, Central Connecticut State University, 2005. Accessed at: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1475-6803.2005.00113.x

Doyle, Timothy M. *"Ratings that don't rate; The subjective world of ESG rating agencies."* Research paper, American council for capital information, 2018. Accessed at: http://accfcorpgov.org/wp-content/uploads/2018/07/ACCF_RatingsESGReport.pdf

Tamimi, Nabil & Sebastianelli, Rose. *"Transparency among S&P 500 companies: an analysis of ESG disclosure scores."* Research paper, University of Scranton, 2017. Accessed at: https://www.emeraldinsight.com/doi/10.1108/MD-01-2017-0018

Renneboog, Luc, Horst, Jenke ter & Zhang, Chendi. *«Socially responsible investments: Methodology, Risk, exposure and performance."* Research paper, Tilburg University, TIAS School of Business and Society & University of Warwick, 2007. Accessed at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=985267

Auer, R. Benjamin & Schuhmacher, Frank. "*Do socially (ir)responsible investments pay? New evidence from international ESG data"* Research paper, University of Leipzig, 2015. Accessed at: https://www.sciencedirect.com/science/article/pii/S1062976915000770?via%3Dihub

Thomson Reuters Eikon. (2018) *"ESG scores methodology"*. Accessed at: http://zeerovery.nl/blogfiles/esg-scores-methodology.pdf

Forbes. (2017) *"Socially responsible investing earn better returns from good companies"* Accessed at: https://www.forbes.com/sites/moneyshow/2017/08/16/socially-responsible-investing-earn-better-returns-from-good-companies/?fbclid=IwAR2o4cMX5NfSpX8a4IgpXNtUbNVzrC50eT-zdPHSMza8PH0QewsHkhkwnkc#14a5a5d4623d

Federal department of Foreign Affairs Switzerland, The Global Compact & International Finance Corporation. (2005). *"Who cares wins"*. Accessed at: https://www.ifc.org/wps/wcm/connect/9e4b228048855ccf8cd4de6a6515bb18/WhoCaresWins_2005ConferenceReport.pdf?MOD=AJPERES&CACHEID=9e4b228048855ccf8cd4de6a6515bb18

Thomson Reuters Eikon. *"Thomson Reuters Business Classification."* Refinitiv. Accessed at: https://www.refinitiv.com/content/dam/marketing/en_us/documents/quick-reference-guides/trbc-business-classification-quick-guide.pdf

CFA Curriculum, Volume 4 (2019). *"Subject 7. Optimal portfolio"* Accessed at: https://analystnotes.com/cfa-study-notes-discuss-the-selection-of-an-optimal-portfolio-given-an-investors-utility-or-risk-aversion-and-the-capital-allocation-line.html

Jed Emerson. *"Risk return and impact: Understanding diversification and performance within an impact investing portfolio".* Issue brief, Impactassets. Accessed at: https://www.impactassets.org/files/downloads/ImpactAssets_IssueBriefs_2.pdf

Fama & French official website. Accessed at: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/index.html

Ackert, Lucy F & Deaves, Richard. (2009) *"Behavioral Finance, Psychology, Decision-making and Markets, international edition"* South-Western Cengage learning

# 9.0 Appendixes

## Appendix 1 – Stata transcript regression year 2009-2013

```
5 year 2009 - 2013

Single Index
 xtreg Return ExcessStoxxEurope ESG


Random-effects GLS regression                 Number of obs     =     55,679 Group

variable: group                         Number of groups  =        928


R-sq:                                         Obs per group:
     within  = 0.0749                                       min =         59

between = 0.0095                                       avg =       60.0

overall = 0.0723                                       max =         60



                                              Wald chi2(2)      =    4441.81

corr(u_i, X)   = 0 (assumed)                  Prob > chi2       =     0.0000


-----------------------------------------------------------------------------------
           Return |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
------------------+----------------------------------------------------------------
ExcessStoxxEurope |   1.090002   .0163712    66.58   0.000     1.057915    1.122089
              ESG |  -.0002184   .0000735    -2.97   0.003    -.0003625   -.0000743
            _cons |   .0518756   .0046542    11.15   0.000     .0427536    .0609976 -

-----------------+----------------------------------------------------------------

sigma_u |  .02703717          sigma_e |  .16478623              rho |  .02621464

(fraction of variance due to u_i)


Fama-French YO

Random-effects GLS regression                 Number of obs     =     55,679 Group

variable: group                         Number of groups  =        928

R-sq:                                         Obs per group:
     within  = 0.0854                                       min =         59

between = 0.0095                                       avg =       60.0

overall = 0.0824                                       max =         60


                                              Wald chi2(5)      =    5119.05
corr(u_i, X)   = 0 (assumed)                  Prob > chi2       =     0.0000

-----------------------------------------------------------------------------------
           Return |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
------------------+----------------------------------------------------------------
ExcessStoxxEurope |   1.036187   .0233992    44.28   0.000     .9903251    1.082048
              ESG |  -.0002184   .0000735    -2.97   0.003    -.0003625   -.0000743
              SMB |   .6311881   .0388999    16.23   0.000     .5549457    .7074306
```

```
             HML |   .1513644   .0345589     4.38   0.000     .0836301    .2190987
             MOM |  -.1912366     .01679   -11.39   0.000    -.2241443   -.1583289
           _cons |   .0500047   .0046613    10.73   0.000     .0408687    .0591408 -

-----------------+----------------------------------------------------------------

sigma_u |  .02713128           sigma_e |  .16385554              rho |  .02668527

(fraction of variance due to u_i)
```

---

```
With Dummy Variables

Random-effects GLS regression                   Number of obs     =     55,679 Group

variable: group                           Number of groups  =        928


R-sq:                                           Obs per group:
     within  = 0.0854                                         min =         59

between = 0.0244                                        avg =       60.0

overall = 0.0830                                        max =         60


                                                Wald chi2(8)      =    5133.25

corr(u_i, X)   = 0 (assumed)                    Prob > chi2       =     0.0000


----------------------------------------------------------------------------------
          Return |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----------------+----------------------------------------------------------------
ExcessStoxxEurope |   1.036187   .0233992    44.28   0.000     .9903251    1.082048
             ESG |  -.0002414    .000074    -3.26   0.001    -.0003864   -.0000964
             SMB |   .6311885   .0388999    16.23   0.000     .5549461     .707431
             HML |   .1513643   .0345589     4.38   0.000       .08363    .2190986
             MOM |  -.1912366     .01679   -11.39   0.000    -.2241443   -.1583289
          France |   .0084228   .0038313     2.20   0.028     .0009137     .015932
           Italy |  -.0130711   .0052995    -2.47   0.014    -.0234579   -.0026843
Energy |  -.0068236   .0046777    -1.46   0.145    -.0159916    .0023445
           _cons |   .0516263   .0046554    11.09   0.000     .0425018    .0607507 -

-----------------+----------------------------------------------------------------

sigma_u |  .02687131           sigma_e |  .16385554              rho |  .02618966

(fraction of variance due to u_i)

-----------------------------------------------------------------------------
```

## Appendix 2 – Test for pooled OLS

```
. xtreg Return ExcessStoxxEurope ESG

Random-effects GLS regression                   Number of obs     =    111,359

Group variable: group                           Number of groups  =        928

R-sq:                                           Obs per group:
     within  = 0.0700                                         min =        119

between = 0.0079                                              avg =      120.0

overall = 0.0685                                        max =        120
```

```
                                                Wald chi2(2)      =    8315.67

corr(u_i, X)   = 0 (assumed)                    Prob > chi2       =     0.0000

------------------------------------------------------------------------------------
           Return |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----------------+------------------------------------------------------------------
ExcessStoxxEurope |   .9345575    .010253    91.15   0.000     .9144621     .954653
              ESG |  -.0001228   .0000453    -2.71   0.007    -.0002116    -.000034
            _cons |   .0293922   .0028641    10.26   0.000     .0237786    .0350059 -

-----------------+-----------------------------------------------------------------

sigma_u |   .0173105          sigma_e |  .13406079              rho |  .01639964

(fraction of variance due to u_i)

------------------------------------------------------------------------------------

. xttest0

Breusch and Pagan Lagrangian multiplier test for random effects


        Return[group,t] = Xb + u[group] + e[group,t]

        Estimated results:
                                 |       Var     sd = sqrt(Var)
                        ---------+-----------------------------

-                   Return |   .0196149       .1400534

e |   .0179723       .1340608                        u

|   .0002997       .0173105         Test:   Var(u) = 0

chibar2(01) =  1765.19                           Prob

> chibar2 =   0.0000
```