

UNIVERSITY OF VAASA
SCHOOL OF ACCOUNTING AND FINANCE

Alvar Enala

**THE INCORPORATION OF ESG CRITERIA AND
THE STOCK PERFORMANCE OF FINANCIAL SECTOR**

Master's Thesis in
Finance

Master's Degree
Programme in Finance

VAASA 2019

TABLE OF CONTENTS	page
LIST OF FIGURES, TABLES, AND EQUATIONS	5
ABSTRACT	7
1. INTRODUCTION	9
1.1. Purpose of the thesis and contribution	10
1.2. Research question and hypotheses	11
1.3. Structure of the thesis	13
2. THEORETICAL BACKGROUND	15
2.1. EMH, MPT, and their link to SRI	15
2.2. Socially Responsible Investing	25
2.2.1. History of SRI	28
2.2.2. The position of SRI today	29
2.2.3. SRI strategies	32
2.3. Corporate Social Responsibility	34
2.3.1. Theories of CSR	36
2.4. Literature review	42
3. DATA AND METHODOLOGY	47
3.1. Data sources and description	47
3.2. Portfolio construction and descriptive statistics	54
3.3. Methodology and performance measurement	60
3.3.1. Capital Asset Pricing Model	60
3.3.2. Fama and French multi-factor models	61
3.3.3. Performance measurement	62
4. EMPIRICAL RESULTS AND ANALYSIS	63
4.1. Whole sample period	65
4.2. Post-crisis sample period	71
4.3. Summary of the results	74

5. CONCLUSION	81
LIST OF REFERENCES	85

LIST OF FIGURES, TABLES, AND EQUATIONS

Figure 1. Investment opportunities described by the CAPM	17
Figure 2. The portfolio selection process	21
Figure 3. The growth of PRI between 2006 and 2019	27
Figure 4. Socially responsible investing in the US between 1995 and 2018	30
Figure 5. The distribution of ESG dimensions incorporated by asset managers	31
Figure 6. The Thomson Reuters ASSET4 ESG measures and score formation	48
Figure 7. ESG score distribution over the whole sample period	49
Figure 8. ESG score distribution over the post-crisis sample period	52
Figure 9. The overall return distribution of all the created portfolios	59
Table 1. Descriptive statistics of the ESG scores (whole sample)	51
Table 2. Descriptive statistics of the ESG scores (post-crisis sample)	53
Table 3. Descriptive statistics of the annual excess returns (whole sample)	56
Table 4. Descriptive statistics of the annual excess returns (post-crisis sample)	58
Table 5. The OLS regression results (whole sample, CAPM & 3-factor model)	65
Table 6. The OLS regression results (whole sample, 5-factor model)	68
Table 7. The OLS regression results (post-crisis, CAPM & 3-factor model)	71
Table 8. The OLS regression results (post-crisis, 5-factor model)	73
Table 9. Summary of the OLS regression results (whole sample, 2002-2017)	75
Table 10. Summary of the OLS regression results (post-crisis, 2010-2017)	76
Table 11. Summary of the hypotheses and evidence	80
Equation 1. Capital Asset Pricing Model (CAPM)	18
Equation 2. Fama and French (1993) three-factor model	19
Equation 3. Fama and French (2015) five-factor model	19
Equation 4. Fama and French (2018) six-factor model	20
Equation 5. Basic utility function	24
Equation 6. Utility function for an ethical investment	24
Equation 7. Holding period return (HPR)	55

UNIVERSITY OF VAASA
School of Accounting and Finance

Author:	Alvar Enala	
Topic of the Thesis:	The Incorporation of ESG Criteria and the Stock Performance of Financial Sector	
Degree:	Master of Science in Economics and Business Administration	
Master's Programme:	Master's Degree Programme in Finance	
Name of the Supervisor:	Timo Rothovius	
Year of Entering the University:	2015	
Year of Completing the Thesis:	2019	Pages: 90

ABSTRACT

A substantial amount of academic research argue that there is an anomaly between high ESG scored companies and financial performance, yet many academics also reject such a view. Thus, this thesis investigates whether the incorporation of high ESG criteria leads to abnormal stock performance in the financial sector. The thesis uses ESG and financial data obtained from the Thomson Reuters ASSET4 database and factor data obtained from Kenneth R. French's web page. The data sample consists of 193 financial companies, divided into banks and financial services companies, that are or were listed in the NYSE between 2002 and 2017. Financial sector is of interest due to its troublesome reputation in corporate responsibility and ethics altogether.

The thesis employs three different asset pricing models in order to minimize the impression of p-hacking and thereby investigate whether financial companies with higher ESG scores overperform, underperform, or does neither when compared to a risk-free investment. These are asset pricing models are the CAPM, the Fama and French (1993) three-factor model, and the Fama and French (2015) five-factor model. The regression analyses is divided into two separate sample periods: the first sample covers the whole data period, whereas the post-crisis period is covering the more recent years during which the ESG phenomenon has emerged significantly. The thesis makes use of the best-in-class (worst-in-class) approach by screening 20% of the best (worst) ranked financial companies by their individual or combined ESG criteria.

The thesis finds evidence that incorporating higher ESG criteria leads to either negative abnormal stock returns or does not have an effect at all. According to the OLS regression results obtained with the Fama and French three-factor model, the top financial companies ranked by their "Environmental" scores, generate annual alpha of -5.94% over the whole sample, whereas the top financial companies ranked by their "Governance" scores generate annual alpha of -4.66%. This indicates that the high ESG scored financial companies underperform the risk-free investment. Furthermore, the CAPM and the Fama and French five-factor models provide statistically insignificant alphas, therefore indicating that high ESG scored financial companies neither under- or overperform.

KEYWORDS: ESG, CSR, Responsibility, Financial Sector, Alpha

1. INTRODUCTION

“The issue of whether companies should consider their social responsibility or the impact of their activities on their stakeholders is no longer up for discussion. These issues ... have become a central part of the creation of shareholder value and the management of both global and local enterprises.”

As stated by Epstein (2018), the question whether companies should take their corporate social responsibility (henceforth “CSR”) or the impact of their other operations into consideration has become self-evident. The question has moved from “whether” to “how” to incorporate Environmental, Social, and Governance (ESG) criteria into the everyday decision-making process of management. The issues today are, how to be more sustainable and socially responsible, how to engage stakeholders more effectively, and what are the specific actions that executives should implement in order to deal with the contradiction of trying to improve CSR and financial performance at the same time. (Epstein 2018: 19.)

Conventional assets pricing models such as the Capital Asset Pricing Model (henceforth “CAPM”), described for instance by Sharpe (1964), Lintner (1965), and Mossin (1966), share the collective assumption that all rational investors evaluate the risk-return tradeoff of their investments when making investment decisions. However, there are many examples in the academic literature that indicate violations to this assumption. Socially responsible investing (henceforth “SRI”) is one of the more recent and well-known violation to the assumption that investors care only about the payoff of their investment. The goal of SRI is not necessarily to maximize one’s profits, but rather to give a statement that an investment is in line with one’s personal, political, religious, environmental, social, and/or ethical concerns. (Beal, Goyen, & Phillips 2005.)

Socially responsible investing has experienced a remarkable growth phase during the last two decades, of which the growth has been even more remarkable over the last few years. The industry of sustainable and responsible investing has grown 18-fold since 1995, and in addition, it has matured and expanded across numerous asset classes over the same

period. In the United States alone, SRI assets have expanded to \$12.0 trillion, meaning a significant growth rate of 38% from \$8.7 trillion in 2016. Furthermore, the SRI assets are held by 496 institutional investors, 365 money managers, and 1,145 community investing financial institutions. (US SIF 2018.)

The total amount of assets under management (AUM) of registered investment companies incorporating ESG criteria continues to grow rapidly. From 2016 to 2018, ESG assets in mutual funds expanded to \$2.6 trillion, meaning a significant growth rate of 34%. Also, the number of exchange traded funds (ETFs) more than doubled from 25 to 69. Moreover, The ESG AUM of alternative investment vehicles has tripled since 2016, reaching to \$588 billion at the beginning of 2018. This represents a total of 780 alternative investment vehicles, including hedge funds, private equity & venture capital funds, real estate investments trusts (REITs), and other property funds. Lastly, the community investing sector, including credit unions, community development banks, as well as loan & venture funds experienced a growth rate of 50% between 2016 and 2018, hereby ESG assets reaching to \$185.4 billion. (US SIF 2018.)

1.1. Purpose of the thesis and contribution

Purpose

The purpose of this thesis is to examine whether higher ESG scores lead to abnormal stock returns in the financial sector. In this thesis the financial sector is divided into two subcategories: banks and financial services, thus excluding for example insurance companies, private equity firms as well as REITs. The thesis will be focused only on the financial markets of United States, more specifically on the banks and financial services companies listed in the New York Stock Exchange (henceforth “NYSE”) over a sample period between January 2002 and January 2017. Years 2018 and 2019 are not included in the data sample of this thesis, as the ESG data was just partially available or not available at all.

Contribution

Firstly, this thesis employs new and rarely used ESG data as well as some new methods to re-investigate the previously established academic literature about the relationship between high ESG scores and stock performance. This will contribute more clarity to the question whether financial companies with high ESG scores overperform, underperform, or does neither compared to the risk-free investment. Secondly, this thesis partially models the methods used for instance by Derwall, Guenster, Bauer, and Koedijk (2005), Kempf and Osthoff (2007), Renneboog, Ter Horst, and Zhang (2008), as well as by Halbritter and Dorfleitner (2015) among others. Hereby, evaluating the results conducted by these authors as well as possibly providing an additional narrower answer regarding the research question mentioned later on. All the previously mentioned authors make use of the Capital Assets Pricing Model, Fama and French (1993) three-factor model, or the Carhart (1997) four-factor model, which all are prior versions of the Fama and French (2015) five-factor model that is used in the empirical part of this thesis among with the CAPM and three-factor model. Moreover, what comes to the data and methods, both Thomson Reuters ASSET4 database as well as Fama and French (2015) five-factor model are very rarely used in the academic literature of ESG, CSR, or SRI. Hereby, this thesis will also evaluate the usability of them both in the environment of corporate social responsibility and socially responsible investing.

1.2. Research question and hypotheses

Once again, the purpose of this thesis is to examine whether higher ESG scores lead to abnormal stock returns in the financial sector. For a background, Auer (2016) quite recently showed that it is still possible to achieve abnormal stock returns with SRI in Europe with portfolios screened on the highest ESG scores, i.e. with best-in-class portfolios. However, Renneboog et al. (2008) find that investors who are employing ESG as an investment criteria tolerate higher costs than traditional investors and are thus willing to accept inferior financial performance. Hereby, the research question of this thesis is put as follows:

RQ: “How does the incorporation of high ESG criteria affect stock performance in the financial sector?”

In order to answer this research question, the following three hypotheses are established:

H₀: Incorporating high ESG criteria leads to neither positive nor negative abnormal stock returns in the financial sector.

The null hypothesis (H₀) is the hypothesis that is trying to be rejected. This null hypothesis will *hold* if statistical significance does not exist in the data sample, that is, in the set of given observations. More specifically, there might occur some abnormal stock returns, i.e. alpha (α), when creating portfolios according to their high ESG criteria, yet the probability value of the statistical model indicates that the results are not statistically significant. Furthermore, the null hypothesis will be *rejected* if one of the alternative hypotheses, H₁ or H₂, are proven to be true. In other words, if incorporating high ESG criteria, in fact, does lead to positive or negative abnormal stock returns in the financial sector.

Furthermore, one of the more recent studies of Belghitar, Clark, and Deshmukh (2014) suggest that there is no significant difference between the performance of socially responsible investments and conventional investments. By using previous research and empirical mean-variance evidence, the authors find the results to be truly insignificant. Some other previous studies that find the same insignificance are from Hamilton, Jo, and Statman (1993) as well as from Bauer, Koedijk, and Otten (2005). These studies will be gone through in-depth later in the “Literature review” chapter.

H₁: Incorporating high ESG criteria leads to positive abnormal stock returns in the financial sector.

This hypothesis complies with the study of Auer (2016), as he quite recently showed that it is still possible to achieve abnormal stock returns with SRI in Europe with portfolios screened on the highest ESG scores. Partially modeling Auer, this thesis uses a new data

set and methods to re-examine whether higher ESG scored financial companies can achieve abnormal stock returns in the United States. Also earlier academic research, such as Kempf's and Osthoff's (2007), state that the incorporation of SRI can lead to high positive abnormal stock returns as simply as following a long-short strategy, positive screening approach, or best-in-class approach. According to their research the best-in-class approach can lead to abnormal returns as high as +8.7% per annum.

H₂: Incorporating high ESG criteria leads to negative abnormal stock returns in the financial sector.

This hypothesis adapts with the study of Renneboog et al. (2008), who find in their paper that investors who are using ESG as an investment criteria tolerate higher costs and are thus willing to accept suboptimal financial performance. Also, Halbritter and Dorfleitner (2015) demonstrate that ESG portfolios do not yield any abnormal returns when comparing companies with high and low ESG ratings.

1.3. Structure of the thesis

The first chapter of this thesis is an introduction chapter that conducts the topic, background, motivation, purpose, contribution, research questions and hypotheses, as well as the structure of the thesis. In the second chapter, the theoretical background will be introduced, more specifically the efficient market hypothesis (EMH), modern portfolio theory (MPT), and the theories behind socially responsible investing (SRI). The aim is to describe the main theoretical framework behind these theories, but also to examine the link between EMH, MPT and SRI. Furthermore, the second chapter continues by focusing more carefully on socially responsible investing. This chapter is divided into three parts: history of SRI, present-day of SRI, and SRI strategies. The latter part of the second chapter continues with describing the framework for corporate social responsibility, and the subchapter 2.3.1. describes the theories behind CSR in-depth. The second chapter will also concentrate on prior empirical evidence in a form of literature review. The review is divided into three subcategories, categorized by the effect SRI has on stock markets: positive, negative, and insignificant effect.

After the first two chapters of introduction and theoretical background, the third chapter focuses on the data and methods used in the thesis. Chapter 3.1. provides a detailed description of the data and its sources. After that, chapter 3.2. introduces the portfolio creation process in a great detail and with examples. Lastly, the end of the chapter 3. will focus on the methodology and performance measurement used in the thesis. Furthermore, chapter 4. shows and explains the obtained results of the regression analyses as well as explains in detail how they relate to the research question and hypotheses. This chapter is divided into three parts: results on the whole sample period, results on the post-crisis sample period, and summary of the results. In the end, chapter 5 concludes the thesis by summarizing all the results as well as presenting limitations and some possible further research on the topic.

2. THEORETICAL BACKGROUND

This chapter aims to provide the main theoretical framework behind the thesis in order the reader to comprehensively understand the complete proportions of the study. The first part of the chapter presents the efficient market hypothesis as well as the modern portfolio theory, thereafter following the theory behind socially responsible investing. What comes to the first two theories, they are fundamental, since SRI can negatively affect market efficiency. All in all, the chapter aims to connect EMH and MPT to SRI with the assistance of current literature and academic research. Furthermore, this chapter also aims to provide a much broader and detailed perspective on socially responsible investing as well as on corporate social responsibility. Chapter 2.2. provides a thorough review on socially responsible investing with subsections revolving around the history, present status, and the main strategies of SRI. Furthermore, chapter 2.3. offers a detailed review on corporate social responsibility with the main focus on the theories behind it. The chapter is then concluded with a literature review on the topic.

2.1. EMH, MPT, and their link to SRI

Efficient Market Hypothesis (EMH)

The theory of efficient capital markets was originally presented by Eugene Fama. According to Fama (1970), capital markets are generally referred to as efficient when stock prices fully reflect all the available information. In efficient markets, securities are traded at their intrinsic value, meaning that they are neither over- or undervalued and hereby arbitrage opportunities do not occur. However, this assumption of fully efficient markets does not always hold, thus three different forms of market efficiency are presented. (Fama 1970.) The next paragraph presents these different forms of market efficiency, which are: the strong form, semi-strong form, and weak form.

When markets are in a weak form of efficiency, security prices reflect only historical information. In other words, when the terms of weak form efficiency are satisfied, it is

impossible to make any abnormal returns based on the historical information of security prices. This means that the weak form of market efficiency is based on the “random walk” theory, stating that securities’ market movements are random, making it impossible to anticipate future stock prices. Furthermore, in a semi-strong form of market efficiency, all the historical stock information as well as public information are reflected into stock prices. Public information is considered to be for example profit predictions, dividends, announcements of mergers and acquisitions, and other corporate actions. Finally, a strong form of market efficiency is achieved when stock prices reflect all the available information, including information that is not public, as well as historical (weak form) and public (semi-strong form) price information. This non-public information can be for example referred to as insider information, since some stakeholders, for instance CEOs, have monopolistic access to important information regarding stock price movements. (Fama 1970.)

Subsequently, Fama has revised his original theory of efficient capital markets. In his second main publication of efficient capital markets, Fama (1991) develops his original theory by reviewing the empirical and theoretical research behind the EMH. Furthermore, he also alters the different forms of capital market efficiency. First, the weak form of market efficiency, which so far comprises only the forecasting power of past stock returns, is now revised to include a more general range of test for the predictability of returns. Thus, in addition to the forecasting power, these return predictability tests also contain variables such as interest rates and dividend yields. These tests also consist returns’ cross-sectional predictability as market efficiency and equilibrium-pricing are indivisible from one another. Furthermore, asset pricing models are also included in the tests as well as anomalies such as the January effect and size factor. Continuing, for the semi-strong and strong forms of market efficiency, Fama (1991) proposes a title change, however the content remaining the same. The previous title of semi-strong form is now changed to “event studies” because the use of event studies is increasing as the gathered evidence on market efficiency from them are the most supportive and direct. Finally, the title of strong form is altered to “tests for private information”, as the name describes the notion of insider information much more distinctively. (Fama 1991.)

Continuing, the CAPM of Sharpe (1964) and Lintner (1965) signifies the birth of asset pricing theory. It provides a strong and instinctively pleasing predictions about how to measure the relationship between expected return and risk as well as risk itself. Moreover, the CAPM builds around the portfolio selection theory introduced by Markowitz (1952). According to the model, investors select a portfolio at a past time $t-1$ that produces an unpredictable return at a future time t . This model hypothesizes that investors are risk averse and that they care only about the variance and mean of their investment when choosing among investment portfolios. Thus, as a consequence, investors select “mean-variance-efficient” portfolios that i) maximize the expected return with the given variance, and ii) minimize the variance of portfolio return with the given expected return. (Markowitz 1952; Fama & French 2004.)

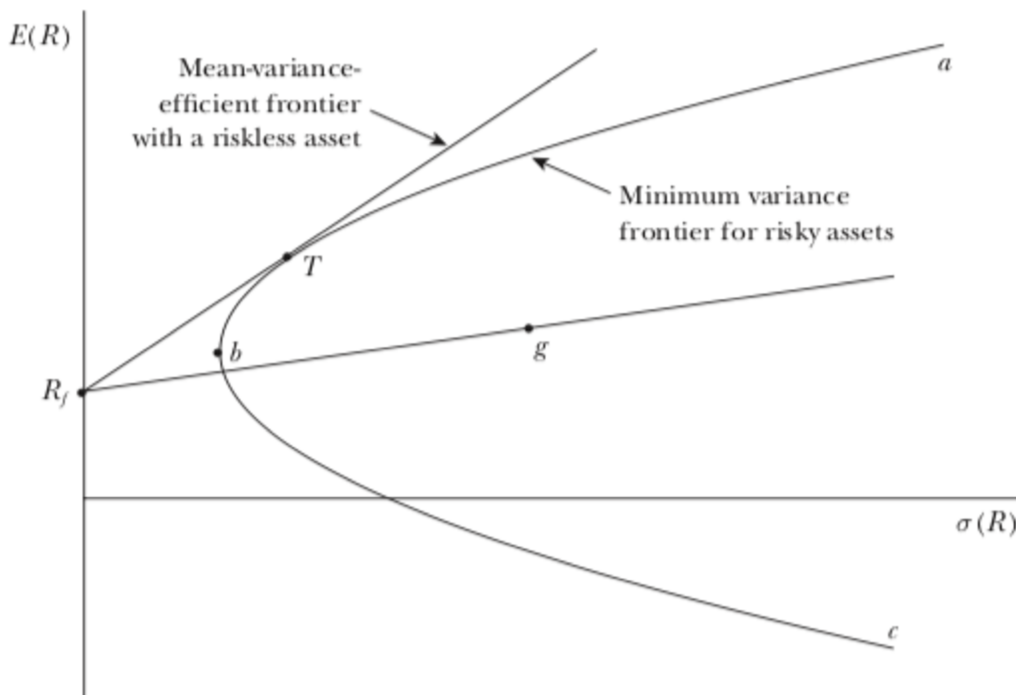


Figure 1.) Investment opportunities described by the CAPM (Fama et al. 2004).

The above figure illustrates the CAPM portfolio opportunities: the vertical axis presents the expected return, $E(R)$, and the horizontal axis presents the portfolio risk, $\sigma(R)$, measured by the standard deviation of portfolio return. The curve abc illustrates the minimum variance frontier for risky assets, meaning it tracks combinations of expected

return and risk that minimize return variance at different levels of expected return. Thus, the trade-off between expected return and risk for minimum variance portfolios is evident: an investor who seeks for a high expected return, at a point a for instance, must accept high volatility. Moreover, at point T , where the minimum variance frontier for risky assets and the mean-variance-efficient frontier for risk-free assets meet, the investor can have a moderate rate of expected return with much lower volatility. (Fama et al. 2004; Bodie, Kane, & Marcus 2014.)

The CAPM is still broadly used in finance because of its comprehensibility. It is used in many financial applications, such as approximating the cost of capital as well as assessing the performance of managed portfolios. The model illustrates the relationship between systematic risk and expected return for a certain asset or a portfolio: as the total risk of an asset or a portfolio increases, investors start to demand higher expected returns for it. (Fama et al. 2004; Bodie et al. 2014.) Hereby, the CAPM formula can be put as the following:

$$(1) \quad R_{it} = R_{Ft} + \beta_i(R_{Mt} - R_{Ft}),$$

where: R_{it} = return on security or portfolio i for period t

R_{Ft} = risk-free rate of return

R_{Mt} = rate of return for a market portfolio

β_i = beta coefficient of an investment, i.e. systematic/market risk.

Even though the CAPM is still broadly utilized in finance, perhaps even better way to measure the capital market efficiency is the Fama and French three-factor model. The Fama and French (1993) three-factor model is designed to describe stock returns, hereby complementing the CAPM. It was established to take into account the factors that are not explained by the CAPM. These are, the excess return on a market portfolio ($R_{MT} - R_{FT}$), the return on a diversified portfolio of small stocks minus the return on a diversified portfolio of big stocks (SMB_t), as well as the difference between the returns on diversified portfolios of high and low book-to-market stocks (HML_t). (Fama & French 1993; 2015.)

$$(2) \quad R_{it} - R_{Ft} = \alpha_i + \beta_{1,i}(R_{Mt} - R_{Ft}) + \beta_{2,i}SMB_t + \beta_{3,i}HML_t + \epsilon_{i,t},$$

where: R_{it} = return on security or portfolio i for period t

R_{Ft} = risk-free rate of return

R_{Mt} = rate of return for a market portfolio

α_i = unexplainable portion of the return, alpha

SMB_t = “Small Minus Big”

HML_t = “High Minus Low”

$\epsilon_{i,t}$ = zero-mean residual, i.e. the error term.

Furthermore, the Fama and French (2015) five-factor model was established to improve the previously mentioned three-factor model by adding two new additional factors in it: RMW_t , or “Robust Minus Weak” and CMA_t , or “Conservative Minus Aggressive”. Hereby, the five-factor model can be put as the following:

$$(3) \quad R_{it} - R_{Ft} = \alpha_i + \beta_{1,i}(R_{Mt} - R_{Ft}) + \beta_{2,i}SMB_t + \beta_{3,i}HML_t + \beta_{4,i}RMW_t + \beta_{5,i}CMA_t + \epsilon_{i,t}.$$

The RMW_t (profitability) factor measures the difference between the returns of a diversified portfolio consisting of stocks with high and low profitability, whereas the CMA_t (investment) factor measures the difference between the returns of a diversified portfolio consisting of companies with low investment rate and high investment rate. According to Fama and French (2015), this five-factor asset pricing model is capable of explaining up to 94% of the cross-section variance of the observed portfolios’ returns. (Fama et al. 2015.)

Lastly, the Fama and French (2018) six-factor model is established in the paper “Choosing Factors” (2018), in which Eugen Fama and Kenneth French develop insights as well as test the usability of earlier asset pricing models. Furthermore, the authors somewhat reluctantly add a momentum factor to the five-factor model even though it lacks clear theoretical justification. Fama and French (2018) argue that the five-factor model, and sometimes even the three-factor model, are sufficient enough and therefore

adding a momentum factor would not be fruitful. For instance, in some cases the size-factor (SMB_t) adds only little utility to the three-factor model, yet makes a great difference in the five-factor model in other cases. Therefore, the authors argue that adding new factors to pricing models lead to significant comparison problems and thus the number of factors should be limited. Nevertheless, by adding the momentum factor, i.e. UMD_t (“Up Minus Down”), the Fama and French (2018) six-factor model can be put as follows:

$$(4) \quad R_{it} - R_{Ft} = \alpha_i + \beta_{1,i}(R_{Mt} - R_{Ft}) + \beta_{2,i}SMB_t + \beta_{3,i}HML_t + \beta_{4,i}RMW_t + \beta_{5,i}CMA_t + \beta_{6,i}UMD_t + \epsilon_{i,t} .$$

In the thesis, the CAPM, the three-factor model, and the five-factor model are used to implement the OLS regression analyses, of which the five-factor model is naturally capable of explaining the largest fraction of the cross-section variance. The six-factor model is not used due to the momentum factor’s imperfection in delivering distinctive theoretical justification for the asset pricing model.

Modern Portfolio Theory (MPT)

This subchapter presents the modern portfolio theory (MPT) that is based on Harry Markowitz’s (1952) study about portfolio selection. According to Markowitz (1952), the portfolio selection process can be divided into two separate stages. The first stage involves the performance analysis of the available securities, starting with observation and experience and ending with beliefs about the future performance. The second stage begins with the appropriate beliefs about future performances, ending with the portfolio selection. (Markowitz 1952.)

Furthermore, a contemporary version of the Markowitz’s portfolio optimization model generalizes the portfolio construction problem as a selection between many risky securities and a risk-free asset. Compared to Markowitz’s (1952) model with two parts, this modern version has three steps in it. The first step includes the determination of available risk-return opportunities to the investor. The minimum-variance frontier of

risky assets summarizes all of these previously mentioned risk-return opportunities. This minimum-variance frontier is a graph, indicating the lowest possible variance that can be obtained for a given expected return of a portfolio. All the individual assets lie on the right hand side of the efficient frontier, at least if short selling is allowed when constructing the risky portfolios. Moreover, the graph indicates that risky portfolios including only one asset, are inefficient, and that the diversification leads to lower standard deviations with higher expected returns. (Bodie et al. 2014: 220.)

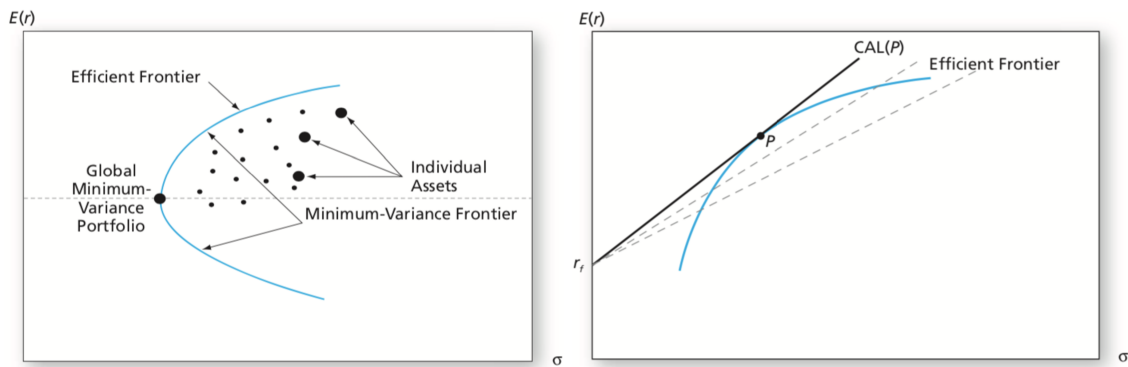


Figure 2.) The portfolio selection process: (left) the minimum-variance frontier of risky assets, (right) the efficient frontier of risky asset with the optimal CAL (Bodie et al. 2014: 220-221).

As can be seen from the figure, all the portfolios that are lying on the minimum-variance frontier, above the graph's dotted line, offer the best risk-return combinations and are hereby candidates as the optimal portfolio. The part of the efficient frontier that is lying above the global minimum-variance portfolio is called the efficient frontier of risky assets, indicating the combinations of expected return and standard deviation. Therefore, portfolios that lie under the global minimum-variance portfolio are considered as inefficient since there is a portfolio with higher expected return and the same standard deviation situated right above it. (Bodie et al. 2014: 220.)

Furthermore, the second step of the optimization process includes the risk-free asset by searching for the capital allocation line (CAL) with the steepest slope (in other words, with the highest reward-to-variability ratio, or Sharpe ratio) as shown in the figure earlier. The CAL, supported by the optimal portfolio P , is tangent to the efficient frontier, meaning that this CAL dominates all the alternative feasible lines (the dotted lines that

are drawn through the efficient frontier). Therefore, meaning that the portfolio P is the optimal risky portfolio. Finally, in the third and last step of the optimization process, investor selects a suitable mix between the optimal risky portfolio P and risk-free assets, such as government bills. (Bodie et al. 2014: 220-221.)

Socially Responsible Investing (SRI)

As already mentioned, the MPT acknowledges that investors can decrease their risk exposure by means of diversification yet without affecting future returns. This “portfolio effect” resulted from the diversification is in fact an outcome caused by the imperfect correlation of returns between securities. The lower the correlations are, the greater the decrease in risk is. Moreover, when a portfolio is well-diversified, it involves only economy-wide risk that cannot be diversified, recognized as market risk. Thus, taking social criteria into account when making investments decisions should theoretically damage the “portfolio effect” by increasing risk exposure and decreasing returns. Hereby, portfolios that are incorporating social criteria should experience lower returns and thereby be suboptimal. However, academic research has shown that aspects of social responsibility, such as the incorporation of social criteria, can in fact enhance the “portfolio effect” by decreasing the overall risk of a portfolio. This is because the funds that are incorporating social criteria provide differing market correlations in comparison to conventional funds, thus providing financial benefits to investors, especially during the times of macroeconomic crisis and market turmoil. (Hickman, Teets, & Kohls 1999.)

Continuing, according to general perception in financial studies, investors are considered to act rationally when prioritizing their expected payoffs and when making decisions on mean-variance optimization. Hereby, investors’ individual motives, values, or preferences are not taken into consideration when examining the behavior of these investors. If investors would behave as the traditional theory in finance supposes, SRI would only exist since it offers the possibility for same returns at a lower level of risk or superior returns at the same level of risk. Over the last two decades, academic research in finance has shifted from the perception of rationality to more psychological ways to explain the phenomena of finance through human behavior. This field of finance is known

as behavioral finance – investors are not considered as rational but rather irrational to a great extent. Examples of these irrationalities are considered to be over-optimism, confirmation bias, conservatism, anchoring, and framing among many others. Furthermore, according to the present assumption in finance, investment decisions are often made on the basis of risk-return tradeoff, meaning that investors demand increasingly higher compensation as the level of risk increases. (Beal et al. 2005.)

According to Beal et al. (2005), there are three possible reasons why investors may favor ethical investment opportunities, which are: superior financial returns, contribution to social change, and/or non-wealth reasons. These investment motives are neither exclusive or exhaustive, yet they offer a functional starting point to understanding and analyzing the behavior of ethical investors. The relationship between the level of corporate social responsibility of a company and its financial performance is a broadly researched topic in the academic literature. Moreover, it is still somewhat uncertain whether SRI funds under- or overperform compared to the conventional investment funds, thus making them legitimate investment opportunities. As mentioned, besides to superior financial returns, socially responsible investors are also driven by non-wealth reasons. This is shown through the investors' willingness to bear additional transaction costs for investment opportunities that are in line with one's values and preferences. In a nutshell, socially responsible investors make their investment decisions based on the real outcomes of the firm's operational activities, as their fundamental goal is to accomplish significant social change through companies that are operating in sustainable and socially responsible industries. (Beal 2005.)

The utility of socially responsible investment

Continuing with the paper of Beal et al. (2005), the authors argue that the theoretical framework for SRI can be controlled from the utility function of an ethical investor. In addition to financial returns, the utility function of the ethical investment also yields a flow of pleasure as well as social status. This utility function comprises of the investor's risk tolerance, expected returns for the risky investment as well as of the utility of investing ethically. The utility that the investor can achieve from taking part in the

investment depends on whether the investment is considered as sustainable, responsible, and ethical, or vice versa. If the investment is conceived to be unethical, the socially responsible investor gains less utility than he would avoiding it. Vice versa, if the investment is considered as ethical, the investor gains more utility than he would avoiding it. In addition, as the size of the investment decreases, the expected utility of the investment increases. (Beal 2005.)

Hereby, positioning SRI within a theoretical framework can be also approached by making slight adjustments to the ethical investor's utility function. In modern finance, the basic utility function explaining investor behavior typically includes two variables: expected return and risk. (Beal et al. 2005.) Thus, the basic utility function can be presented as follows:

$$(5) \quad U = f(E_R, \sigma_R) ,$$

where: U = utility
 E_R = expected return
 σ_R = standard deviation, i.e. volatility.

However, to adjust this particular utility function in order to serve ethical investing, an additional variable is needed to attach into the function. This additional variable (e) can be named as the “degree of ethicalness” of an investment. (Beal et al. 2005.)

$$(6) \quad U = f(E_R, \sigma_R, e) .$$

As the “degree of ethicalness” is attached to the utility function, the traditional investor's indifference curve changes into ethical investor's indifference plane. The indifference curve of a traditional investor is upward sloping, meaning investors expect higher returns as the risk-level of an investment increases. However, the ethical investor's indifference plane takes so called risk-return-ethicalness tradeoff into account. Thus, compared to traditional investors, socially responsible investors are ready to accept decreasing levels

of expected returns, as they are also taking the investment's ethicalness into account in addition to the similar risk-return tradeoff as traditional investors. (Beal et al. 2005.)

2.2. Socially Responsible Investing

Best known as socially responsible investing (SRI), yet also known as socially conscious, sustainable or ethical investing, refers to an investment strategy that aims to conjoin environmentally sustainable and/or social dimensions as well as benefits acquired from good corporate governance with financial returns. Thus, when implementing a strategy of SRI, investors are often ready to sacrifice financial profits in order to achieve better environmental, ethical, or social benefits. In the recent decades, sustainable and responsible investing has emerged as a dynamic and rapidly growing segment in the financial sector of United States. (Schueth 2003; Brzeszczyński & McIntosh 2014.)

Furthermore, Eccles and Viviers (2011) examine an extensive set of academic research in order to investigate the origins and meanings of names used to describe investment strategies that conjoin a set of environmental, social, and corporate governance (ESG) dimensions in the academic literature. The review consists of 190 academic studies, spanning a 34-year time period between 1975 and 2009. According to the paper, three investment strategies are commonly connected with the name "Responsible Investing": cause-based investing, positive screening, and best-in-class. Moreover, the definition may also be connected with ethical egoism. In addition to this, the paper shows that studies associated with deontological ethics are more commonly connected with the name "Ethical Investing", whereas studies linked with the ethics of ambiguity were less commonly associated with the name. The name "Ethical Investing" is more preferred in the United Kingdom, whereas in the United States it seems to be strictly avoided. Moreover, the name is considerably more commonly used in the literature concerning philosophy, ethics, and business ethics than in literature dealing with economics, finance, and investing. In addition to these two, a set of other more ambiguous names also appear in the academic literature, such as environmentally responsible investing, community investing, mission-based investing, faith-based investing, social choice investing, moral

investing, green investing, red investing, etc. (Eccles & Viviers 2011.) This thesis will mainly employ the definition of “Socially Responsible Investing” or “SRI”, yet other definitions such as “Ethical Investing” and “Sustainable Investing” will be occasionally met as well.

The definition of ESG

The term “ESG” is nowadays broadly used by institutional investors and other financial professionals. It is not only referred to environmental, social, or corporate governance dimensions, but to all non-financial fundamentals that can have an effect to companies’ financial performance, such as human resource management as well as labor and employment standards. The interest towards ESG criteria has increased and financial professionals, such as asset managers and institutions, have started to implement them into their investment strategies. This is mainly because of ESG’s possible effect on the investment’s risk and return, yet also because of the desire to make an impact. As a matter of fact, the majority of all publicly traded equities on a global scale are as of today signed by the United Nation’s Principles for Responsible Investment (UNPRI). These principles obligate institutional investors to involve their portfolio companies in implementing ESG criteria as a part of their corporate strategies, and also in encouraging other investment intermediaries to do the same. (Ho 2016.)

The Principles for Responsible Investment

The “Principles for Responsible Investment” (PRI) was founded in 2005 by a group of world’s largest institutional investors and with the support of United Nations. As of today, PRI is the principal proponent for responsible investing. The main function of the PRI is to understand the investment dimensions of ESG factors and to support its worldwide investor signatory network in incorporating these criteria into their investment strategies and ownership decisions. The PRI functions as a long-term advocate of its signatories, of the economies and financial markets, and eventually of the society altogether. It is truly independent, and it encourages investors to use ESG criteria in order to better manage risk and to improve their returns. (UNPRI 2019b.)

The number of signatories as well as the assets under management (AUM) of PRI has grown consistently since it fully started to operate in 2006. As seen in the figure below, the AUM has experienced a significant growth from \$6.5 trillion in 2006 to \$86.3 trillion in 2019. In addition, the number of signatories has increased from 63 institutional investors in 2006 to 2372 signatories in 2019. (UNPRI 2019a.)

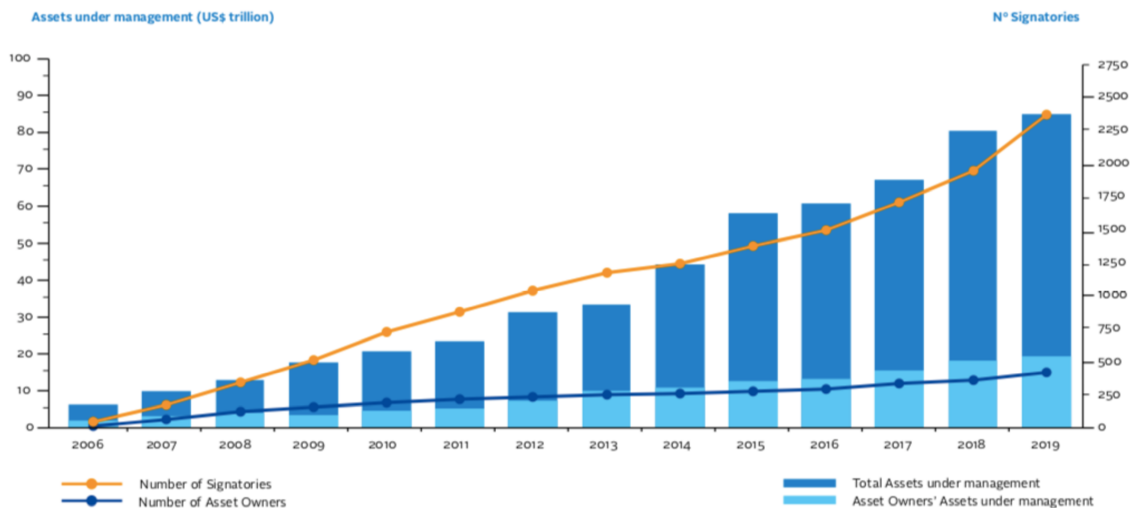


Figure 3.) The growth of PRI between 2006 and 2019 (UNPRI 2019a).

According to the Principles for Responsible Investment (2019a), there are seven principal drivers behind the increasing interest for socially responsible investing. These are: i) the understanding in the financial markets that ESG criteria have a relevant function in determining the risk and return of an investment; ii) the comprehension that incorporating ESG criteria into one's investment strategy is a part of investors' fiduciary responsibility to their beneficiaries; iii) the concern about the negative impact of short-term decision-making on market behavior, firm performance, and investment returns; iv) the legal requirements that are securing the long-term benefits of stakeholders and the financial system altogether; v) the pressure from rival companies that are trying to stand out by offering socially responsible investment opportunities as a competitive advantage; vi) the stakeholders starting to gradually demand for transparency about where their capital is being invested in; and vii) the reputational risks that are destroying firm value in a globalizing world of social media, from concerns such as poor working conditions, employee diversity, pollution, climate change, corruption, etc. (UNPRI 2019a.)

Institutional investors have a responsibility to operate in its stakeholders' best long-term interest. It is shown that in this role ESG factors may have an impact on the performance of the investment portfolios. Thereby, six principles for responsible investment are introduced: 1) incorporating ESG issues into investment analysis and decision-making processes; 2) being active owners and incorporating ESG issues into ownership policies and practices; 3) seeking appropriate disclosure on ESG issues by the entities in which investing in; 4) promoting acceptance and implementation of the Principles within the investment industry; 5) working together to enhance the effectiveness in implementing the Principles; and 6) reporting activities and progressing towards implementing the Principles. (UNPRI 2019a.)

2.2.1. History of SRI

Socially responsible and ethical investing as we know it has ancient origins that date back hundreds of years to Christian, Jewish, and Islamic traditions. Back in the medieval Christian times, there were strict ethical restrictions based on the Bible regarding investments and loans. Also, as commonly known, Judaism has various teachings on how to use money in an ethical way, whereas The Catholic Church instructed an universal prohibition on usury in the 12th century that was relaxed not until the 19th century. Thus, for generations religious investors have avoided investing in companies that profit from unethical products that are meant to enslave or even kill other fellow humans. For instance, it is likely that in the 17th century Quakers introduced the concept of ethical and socially responsible investing to the new world, as they refused to profit from arms trade and slavery. The Methodists, on the other hand, have been using positive screening as their investment criteria for over two hundred years now – they do not want to profit by exploiting others or be involved with any sinful investments. These kinds of investments are nowadays called as “sin stocks”, as they are in the business of tobacco, alcohol, or gambling. Altogether, the deepest religious origins of ethical and responsible investing can still be seen in the United States by the wide-spread avoidance of these previously mentioned “sin stocks”. (Schueth 2003; Renneboog et al. 2008.)

As the ancient ethical investing was based on religious beliefs, the modern version of SRI is rather founded on the individual investor's personal convictions on what is socially acceptable and/or ethical. The modern origins of ethical and socially responsible investing can be traced to the politically fanatic climate of the 1960's. During the restless decade various social campaigns, such as the anti-racism, anti-war, and civil rights movements, have made investors conscious of the social and ethical consequences of their investment decisions. Furthermore, the number of socially conscious investors increased significantly in the 1980's, as millions of people as well as universities, churches, and cities focused on pressuring South Africa's white minority government to close down the racist system of apartheid. Then, with significant amount of new information about ozone depletion and global warming coming out as well as with incidents such as Chernobyl and Exxon Valdez, the investors mindsets moved to even more environmentally conscious and ethical direction. More recently, in the 21st century, human rights issues, such as unhealthy working conditions as well as violent tragedies, such as school shootings, have become focus points for socially responsible investors. (Schueth 2003; Renneboog et al. 2008.)

Altogether, since the early 1990's, the industry of socially responsible and ethical investing has experienced significant levels of growth in the United States and Europe. Ethical consumerism, where consumers pay a considerable premium for products that are in line with their values and ethics, has been an important factor behind this growth. Concerns with human rights, environmental protection, and employment relationships have become common investment screens in socially responsible investing. More recently, frequent corporate scandals have made socially responsible investors more focused on CSR and corporate governance as one of their investment screens. All in all, factors such as sustainability, transparency, and governance have emerged as fundamental criteria for SRI screens. (Renneboog et al. 2008.)

2.2.2. The position of SRI today

The "Report on US Sustainable, Responsible and Impact Investing Trends" (2018) by United States Social Investment Forum (US SIF) argues that the total amount SRI assets

under management (AUM) reached to \$12.0 trillion at the beginning of 2018. This means that one in every four dollars of the total professional AUM in the US are invested in a socially responsible or sustainable way. (US SIF 2018.)

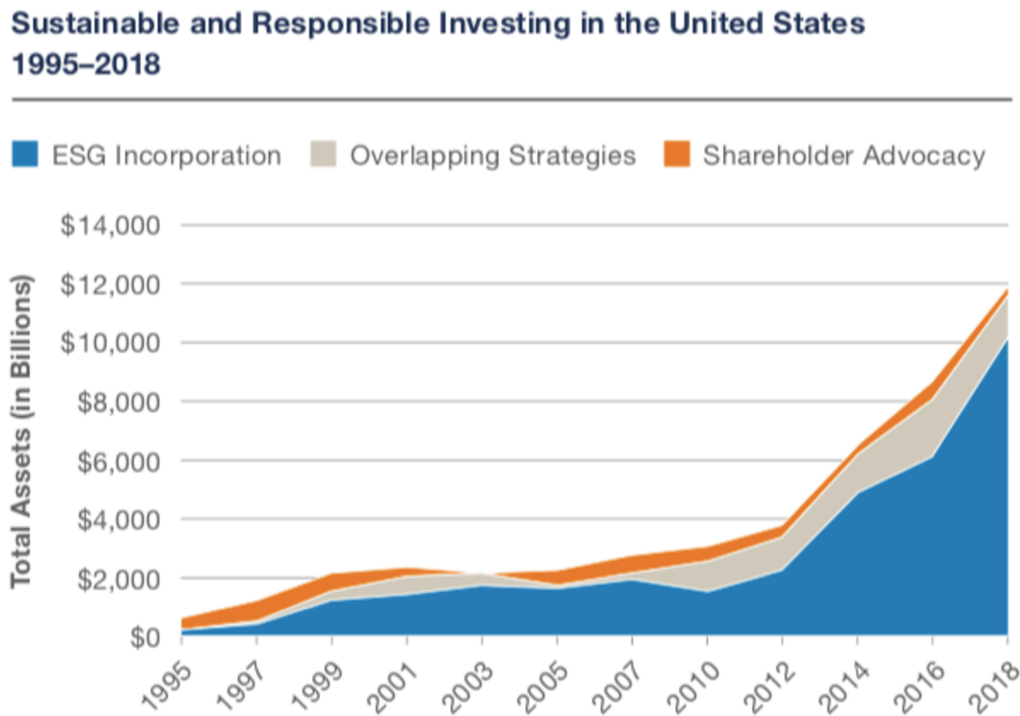


Figure 4.) Socially responsible investing in the United States between 1995 and 2018 (US SIF 2018).

The figure above illustrates the total amount of sustainably and socially responsibly invested assets under management in a time period between 1995 and 2018. Since 1995, these sustainably and responsibly invested assets have grown more than 18-fold, from \$639 billion to \$12.0 trillion, reaching a compounded annual growth rate of 13.6%. When examining more closely, asset managers who applied various different ESG criteria in their portfolio selection and investment analysis, totaled \$11.6 trillion in US-domiciled assets at the beginning of 2018. Moreover, assets managers and institutional investors that filed or co-filed shareholder resolutions between 2016 and 2018 regarding environmental, social, or governance issues towards public companies, totaled \$1.8 trillion in US-domiciled assets at the beginning of 2018. Therefore, after eliminating the overlapping involved in these two different strategies, the total amount SRI assets under management in the US reached to \$12.0 trillion at the beginning of 2018. (US SIF 2018.)

In terms of the individual Environmental, Social, and Governance categories, in 2016 and 2018 asset managers incorporated these ESG dimensions quite evenly across the categories. As the figure below shows, asset managers incorporated social criteria in their investment strategies somewhat more than governance or environmental factors. Moreover, the incorporation of social criteria grew by 39,4% between 2016 and 2018, whereas the incorporation of environmental criteria increased only by 30,2%. (US SIF 2018.)

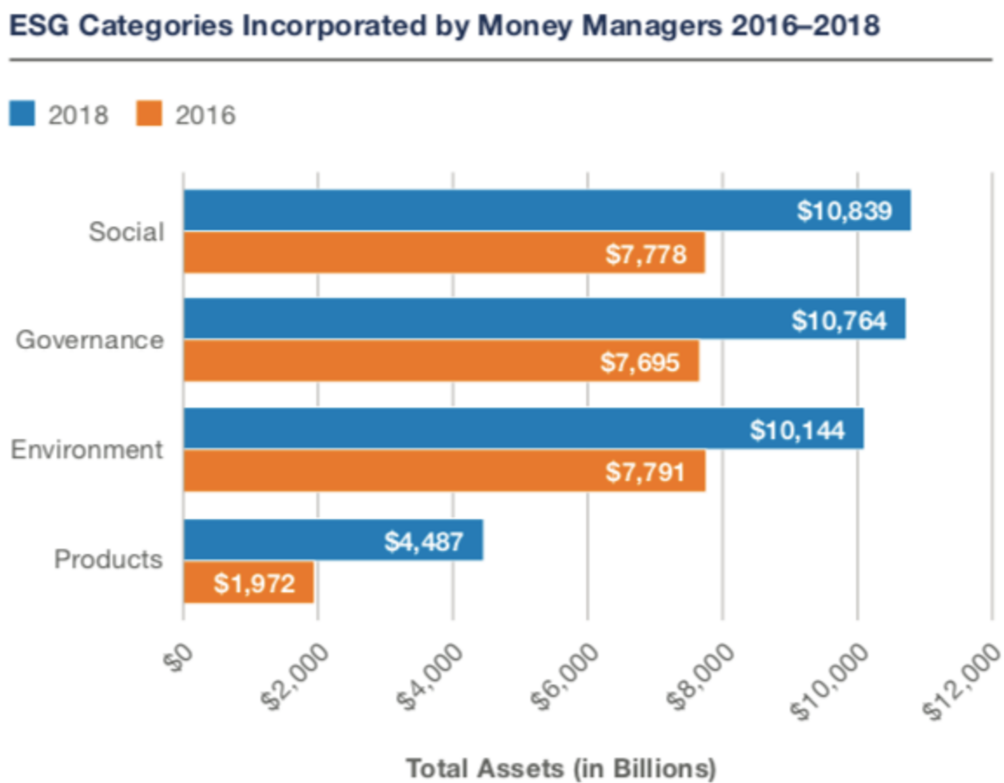


Figure 5.) The distribution of individual ESG dimensions incorporated by asset managers between 2016 and 2018 (US SIF 2018).

Furthermore, between 2016 and 2018 institutional investors incorporated the individual ESG dimensions rather more unevenly across categories. Also institutional investors incorporated social criteria the most, reaching to \$5.2 trillion in 2018, meaning a growth rate of 18,7% from 2016. In comparison, the incorporation of environmental and governance dimensions both reached around \$3.5 trillion in 2018, however both reaching to a growth rate around 40% from 2016 (\$2.5 trillion). (US SIF 2018.)

All in all, as can be concluded from the USSIF's "Report on US Sustainable, Responsible and Impact Investing Trends" (2018), socially responsible investing is currently on the increase and continuing to make a difference. Institutional investors, asset managers, and other financial professionals are demonstrating today a more wide-ranging set of environmental, social, and governance concerns across a wider range of assets compared to few years back. It seems that private individual investors are also gaining a lot of interest in investing in socially responsible and ethical way. Moreover, financial advisors are offering more sustainable and ethical investment opportunities for their clients, as they are also becoming aware of the need and demand for socially responsible investing. Thus, it can be concluded that impactful and ethical as well as socially responsible investing continues to grow rapidly in all areas of investing and finance. (US SIF 2018.)

2.2.3. SRI strategies

This chapter introduces briefly the strategies of socially responsible investing, as understanding them will give the reader more depth about the topic as a whole. According to Schueth (2003), there are three basic strategies intended for the dual-objective of yielding financial returns while at the same time trying to make a difference. These strategies are Screening, Shareholder Advocacy, and Community Investing.

Screening

Screening is the custom of excluding or including firms from investment portfolios based on their environmental, social, and/or governance (ESG) criteria. These investment screens used in socially responsible investing have evolved significantly during the past decades, and today SRI funds typically apply a some sort of combination of these various screens. For instance, SIF's (2003) report shows that more than five screen simultaneously is used by 64% of all the SRI mutual funds in the United States, whereas only one screen is used by 18% of the SRI funds. Moreover, these screens can be categorized distinctly into two separate groups: positive screens and negative screens. (Schueth 2003; Renneboog et al. 2008.)

To begin with, negative screening is presumably the oldest and most rudimentary strategy of SRI there is. It refers to the custom of excluding specific industries or stocks from SRI portfolios based on their ESG criteria. A common negative screen can be applied on an initial asset pool or stock index, such as the NYSE, from which tobacco, alcohol, defense industries, gambling, etc. sin stocks are then excluded. Other negative screens are considered to be for example pornography, abortion, animal testing, and violation of human rights. After conducting the negative screening, portfolios are then constructed through financial and quantitative selection. Secondly, SRI portfolios are today also created via positive screens, which means the selection of specific companies/stocks that satisfy superior standards of CSR. The most typical positive screens are concentrated on the environment, corporate governance, sustainability of investments, and labor relations. Moreover, positive screens are often used with the best-in-class approach, where companies are ranked within each market sector or industry based on their CSR criteria, from which only the best performing companies are selected. (Renneboog et al. 2008.)

Shareholder Advocacy

Shareholder advocacy is considered as the actions socially responsible investors take in their role as responsible shareholders. These actions include participating in a dialogue with companies on their matters of concern as well as voting and submitting proxy resolutions. Socially responsible investors frequently operate together with the management on a course that is believed to improve financial performance over time as well as improve the prosperity of all the corporation's stakeholders – that is, employees, customers, vendors, communities and the natural environment, and especially the shareholders. (Schueth 2003.)

Community Investing

Community investing is a practice of providing capital to low-income people and households, and to communities that are at-risk and have difficulties accessing capital through customary channels. For instance, many socially responsible investors allocate a predetermined percentage of their investments to Community Development Financial

Institutions (CDFIs) with objectives concentrated on offering financing to low-income housing and small business development in underprivileged communities. (Schueth 2003.)

2.3. Corporate Social Responsibility

In the paper “How Corporate Social Responsibility is Defined: an Analysis of 37 Definitions” (2008), Alexander Dahlsrud introduces the two most frequently quoted definitions of CSR as follows:

“A concept whereby companies integrate social and environmental concerns in their business operations and in their interaction with their stakeholders on a voluntary basis.”

“The commitment of business to contribute to sustainable economic development, working with employees, their families, the local community and society at large to improve their quality of life.”

The media, governments, as well as activists have become experts at holding companies accountable for the social and environmental consequences of their operations. Numerous organizations are in the business of ranking companies on the performance of their corporate social responsibility (CSR), and it is attracting a great deal of publicity. As a consequence, many companies have taken CSR as a part of their corporate strategies by making it one of the top priorities. (Porter & Kramer 2006.)

Global climate change, local water and air pollution in Europe, working conditions in the United States, child labor in Asia, and human rights in Africa. These are just a proportion of the challenges that companies and their executives are confronting on a daily basis. The question whether companies should take their social responsibility or the impact of their other operations into consideration has become evident, as these issues have become a fundamental part of the value-creation process for the shareholders and management.

The question has moved from “whether” to “how” to incorporate environmental, social and governance (ESG) dimensions into the everyday decision-making process of the management. In other words, the issues nowadays are, how to be more sustainable and socially responsible, how to engage stakeholders more effectively, as well as what are the specific actions that executives should implement in order to deal with the paradox of trying to improve CSR and financial performance at the same time. (Epstein 2018: 19.)

The academic research on CSR has a propensity to concentrate on the negative examples of bad practice and corporate misbehavior, as unfortunately there are in fact many examples of this. Furthermore, the majority of the corporations that have been labeled as the rogue companies (e.g. oil, tobacco, and chemicals) are currently attempting to distance themselves from the image of corporate misbehavior and bad practice. However, it has been argued that these so-called rogue companies have not in fact been misbehaving any more than the other companies, but have just been caught in their offences. Nonetheless, the distancing of the rogue companies from the others has led to an immense reappearance of interest in corporate behavior that has now been classified as CSR. Thus, companies have been “re-packaging” their corporate activities and behavior as CSR, as there is a lot of evidence that nothing responsible has been actually done but only this “re-packaging”. In other words, it seems that the semiotics are in this case by far more powerful than the actions itself. (Crowther & Seifi 2018: 1.)

The emergence of CSR

The recent emergence of CSR started with the already mentioned rogue companies, them being “the usual suspects”. Due to strong media pressure, governmental regulation, and major disasters, these rogues understood that being involved with human rights issues, supporting oppressive regimes, and polluting the environment, to mention a few, were practices that needed to be changed if they ever wanted to prosper again. Nowadays, however, there is practically no market, industry, or business that has not experienced increasing demand for social responsibility from the society. Furthermore, industries that were for a long time considered as “clean”, such as banking, tourism, entertainment, and retailing, are now facing increasing expectations towards more socially responsible

practices. After the financial crisis of 2007-2009, questions regarding CSR have moved even further to the fore of the political, public, and media interest. The focus has been now mainly on the financial industry, as their careless practices are mostly to blame for initiating the wave of economic recession. (Crane, Matten, & Spence 2019: 3.)

Companies have reacted to this agenda by heavily advocating practices of CSR. Today, companies of all sizes feature CSR reports, departments, managers, or at least projects, and social responsibility is increasingly promoted as a core area of management, next to accounting and finance as well as marketing. As already mentioned in the previous chapter, Crane et al. (2019) also argue that this “new” management style of CSR might in fact just be a way of recycling old practices, and that some of the practices that fall under the CSR label have been relevant for hundreds of years, since the Industrial Revolution. For instance, providing healthcare, ensuring good working conditions, and donating to charity are practices that many industrial companies guaranteed in early 1800’s Europe. Altogether, it can be said that there has been an emergence of so-called CSR “movement”. There has been a significant increase in the number of ESG business consultants, as they are trying to benefit from the phenomena of CSR. Simultaneously, an increasing number of CSR auditors and certifiers are trying to harmonize and institutionalize CSR practices in a global scale. All in all, an increasing number of committed newspapers, websites, and the social media are contributing to provide an identity to CSR as a method of management, and are also helping to establish a global network of CSR practitioners, activists, and academics. (Crane et al. 2019: 3-4.)

2.3.1. Theories of CSR

According to Garriga and Melé (2004), there are four related approaches and main theories of CSR: 1) instrumental theories, in which the company is seen merely as an instrument for financial value-creation and its social activities are only considered as tools to achieve economic results; 2) political theories, which takes the power of companies into account, as businesses have political power in the society; 3) integrative theories, in which companies are focused on satisfying the social demands it experiences; and 4) ethical theories, in which companies have ethical responsibilities that it needs to satisfy.

Next, all of these four CSR theories will be presented in a greater detail, as it is important to fully comprehend them in order to understand to whole depth of corporate social responsibility as a phenomenon.

Instrumental theories

In the group of instrumental theories, CSR is considered only as a strategic tool to accomplish economic objectives – thus, it is ultimately a tool for wealth-creation and profit maximization. As Friedman (1970) famously expresses it: “the only responsibility of business towards society is the maximization of profits to the shareholders within the legal framework and the ethical custom of the country”. Instrumental theories have a long history in finance and they have enjoyed acceptance in business broadly to this day. The concern for profit maximization includes taking into consideration the interest of all stakeholders, not just the interests of shareholders, as it is stated that in some cases the satisfaction of these interests can contribute to maximizing the shareholder value as well. Furthermore, a suitable level of investment in corporate social activities as well as philanthropy is also acceptable for the sake of profits. (Garriga et al. 2004.)

Instrumental theories can be divided into three main groups according to their economic objective. The objective in the first group is to maximize shareholder value, measured by the stock price, which however leads commonly to a short-term profits orientation. In the second group the focus is on the strategic objective of achieving competitive advantages, thus producing long-term profits. In both of these cases, CSR is only a matter of self-interest, as it is only used as an instrument for financial profits. Finally, the third group is connected to cause-related marketing, which means that it is very close to the second group. All in all, several studies have been executed in order to determine the correlation between CSR and corporate financial performance. A growing number of these indicate a positive correlation between CSR and financial performance. Nevertheless, these findings need to be observed with caution, as such correlation is hard to measure in practice. (Garriga et al. 2004.)

Political theories

Political theories of CSR focus on the interactions and connections between the society and businesses as well as on the position and power of businesses and their natural responsibilities. They involve both political analysis as well as political considerations in the discussion of corporate responsibility. Even though there are various different political theories of CSR, according Garriga et al. (2004) to two major approaches can be specified: Corporate Constitutionalism and Corporate Citizenship. (Garriga et al. 2004.)

Davis (1960) was one of the first academics to examine the role of power and the social impact of this power that companies have in the society. By doing so, he introduced the concept of business power as a new element in the discussion of corporate responsibility, and thereby the definition of Corporate Constitutionalism started to gradually born. According to Davis, businesses are social institutions and therefore they have to use their power responsibly. The sources that generate this social power are not merely internal, but external as well. In addition, the business cluster is continuously shifting between the economic, social, and political forums, and thereby it is unstable altogether. Moreover, Davis forms two principles that indicate how social power needs to be managed: “the iron law of responsibility” and “the social power equation”. The iron law of responsibility refers the negative consequences when corporations do not use their social power as intended, whereas the social power equation argues that the social responsibilities of corporations ascend merely from the amount of social power they have.

Even though the idea of companies considered as citizens is not new, a recent interest among practitioners towards this theory has emerged due to particular factors that have had an impact on the relationship between the businesses and society. Four of the most significant factors to mention are the globalization phenomenon and the crisis of the welfare state as well as the decreasing cost with technological improvements and the deregulation process. The two latter have especially led to some multinational companies, such as Google and Facebook, to have greater social and economic power than some governments. The Corporate Citizenship theory seems to give an explanation to this new reality. (Garriga et al. 2004.) Matten, Crane, and Chapple (2003) introduce three outlooks

of Corporate Citizenship: i) a limited view; ii) a view equivalent to CSR; and iii) an extended view of corporate citizenship. Firstly, in the limited view Corporate Citizenship is used in a way close to social investment, corporate philanthropy, or particular responsibilities that are expected from the corporations towards the local communities. Secondly, the view equivalent to CSR, which is rather common, argues that Corporate Citizenship appears as a new conceptualization of the role of corporations in the society, and, this outlook overlaps greatly with the other theories of corporate social responsibility in business and society. Lastly, in the extended view of corporate citizenship, companies are considered in the arena of citizenship at the moment when governments fail to protect the citizenship. This outlook emerges from the fact that some multinational corporations, as mentioned earlier, have today come to replace the most powerful institutions, mainly governments. (Matten et al. 2003; Garriga et al. 2004.)

Integrative theories

Integrative theories of CSR examine how companies integrate social demands, furthermore stating that companies are depended on the society for their existence and continuity as well as growth. These social demands are commonly regarded as the way in which society interacts with companies and gives them prestige and legitimacy. As an outcome, corporate executives should take social demands into consideration and integrate them in a manner that the corporation operates in an agreement with social values. Thus, according to integrative theories, the content of corporate responsibility is constrained to the time and space of each situation depending on the current values of the society. That is, there is not any particular actions that corporations are responsible for conducting through time and in each specific industry. All in all, integrative theories are concentrated on the detection of as well as the response to the social demands that are able to achieve greater social acceptance, legitimacy, and prestige. (Garriga et al. 2004.)

Continuing, integrative theories can be divided into four main approaches, which are: 1) issues management; 2) the principle of public responsibility; 3) stakeholder management; and 4) corporate social performance (CSP). Firstly, issues management refers to social responsiveness, or responsiveness in the face of social issues as well as the process to

manage those issues in the organization. This approach is based on the concept that there is a gap between company's relevant publics expectations of performance and what the company's actual performance is. These gaps are commonly located in a zone called "zone of discretion" in which the firm receives somewhat unclear signals from the environment. Afterwards, the company should distinguish these gaps and close them with a proper response. Secondly, the principle of public responsibility aims to give proper content and substance to help as well as guide the company with its activities of responsibility by limiting the scope of CSR. This school's academics and practitioners criticize the issues management approach (i.e. social responsiveness) as insufficient. They select the term "public" instead of "social" in order to emphasize the importance of the public process over the personal views of morality. The term selection also aims to emphasize that narrow interest groups, such as top executives, do not define the scope of responsibilities, but it is a public process.

Thirdly, instead of focusing on social responsiveness, specific management issues, or on the principle of public responsibility, the stakeholder management approach is oriented, as one would expect, towards stakeholders. Two basic principles of stakeholder management are presented. According to the first, the main goal of stakeholder management is to accomplish the best possible overall co-operation between the whole group of stakeholders and the objectives of the company. The second principle argues that the most efficient strategy to manage stakeholder relations involve efforts that deal with issues affecting multiple interest groups. Finally, the approach of corporate social performance (CSP) involves a search for social legitimacy, with the process of giving proper responses. It is considered to have generally three basic elements: a general definition of social responsibility, a listing of matters in which social responsibility typically appears, and a specific definition of the response philosophy to social issues. Altogether, it aims to give a better picture of the company's corporate social performance with the whole range of responsibilities to society, including the ethical, economic, and legal dimensions of corporate performance. (Garriga et al. 2004.)

Ethical theories

The fourth and last group of CSR theories is considered to be ethical theories, as they concentrate on the ethical conditions that consolidate the relationship between businesses and society. These ethical conditions are founded on principles that indicate “the right thing to do” or the necessity to accomplish a great society. Four main approaches can be distinguished, which are: 1) normative stakeholder theory; 2) universal rights; 3) sustainable development; and 4) the common good approach. (Garriga et al. 2004.)

Stakeholder management approach is occasionally included in the category of integrative theories, as some academics think that this way of management is a method of integrating social requirements. However, today stakeholder management has become an ethical approach mainly because of Freeman’s book in 1984. All in all, this modern normative stakeholder approach that is based on ethical theories, expresses merely a different perspective on CSR, in which ethicality is in central. Continuing, human rights are considered as the foundation for CSR, particularly in the global markets. At the turn of the millennium, some human rights based approaches for CSR were suggested for the first time. One and possibly first of them was presented in 1999 – the United Nations Global Compact, which involves nearly dozen principles in the fields of human rights, environment, and labor. Many other similar principles have been presented later as well. Overall, even though universal rights are a simple question of mutual agreement for many of us, they have a clear theoretical background, and some philosophical theories of moral give them endorsement as well. The third value-based approach, which has become very popular nowadays, is sustainable development. Even though it was originally evolved at a macro-level rather than corporate-level, it requires a significant corporate contribution. There are numerous definitions of sustainable development, however a content analysis of the main definitions propose that sustainable development is “a process of achieving human development in an inclusive, connected, equitable, prudent and secure manner” (Gladwin, Kennelly, & Krause 1995). Lastly, the fourth and final approach, which is less consolidated than the stakeholder theory but still has potential, is the common good approach. The approach of “common good” is a classical idea established in Aristotelian tradition, developed by philosophers, and presumed into today’s social view as a principal

reference for business ethics. According to the approach, businesses, as any other individuals in a society, have to participate to the common good since they are essential parts of the society. Companies participate to the common good in several different ways, for instance by providing goods and services as well as by creating wealth, when at the same time appreciating the dignity and fundamental rights of an individual. To some degree, the common good approach has a lot of similarities with the stakeholder approach as well as with sustainable development, however the philosophical background is different. (Garriga et al. 2004.)

2.4. Literature review

As already mentioned, this study will focus on the stock performance of portfolios incorporating ESG criteria among the financial industry of the NYSE. The majority of academic literature on corporate social responsibility and socially responsible investing of the current decade focuses on analyzing whether the implementation of SRI into investment decisions affects financial performance (Revelli & Viviani 2015). However, the results of academic research on SRI are somewhat contradictory, and thus three differing hypotheses regarding the performance of SRI will be presented.

This chapter will focus on the three mentioned hypotheses regarding the performance of SRI as well as some of the main previous academic research on the topic. These three hypotheses are: SRI has a i) positive effect; ii) negative effect; or iii) insignificant effect on stock performance compared to a risk-free investment.

Firstly, the underperformance hypothesis implies that as the investors set limitations on their investment portfolios due to the screening processes typical in SRI investment strategies, they limit their potential of diversification and thus will achieve less satisfactory risk-adjusted returns in comparison to conventional investment strategies. Secondly, the hypothesis of overperformance rests on the argument that great performance in Environmental, Social, and Governance associated matters lead to a number of positive outcomes. Thus, high ESG scores can be considered as a clear sign of

good management skills, which is leading into better financial performance or cost-reduction during crises. Lastly, the third hypothesis argues that SRI strategies neither adds or destroys value, merely because the expected risk-adjusted returns of SRI stocks are equal to those of so-called traditional stocks. (Renneboog et al. 2008; Revelli et al. 2015.)

Earlier academic research, such as Kempf's and Osthoff's (2007), state that the incorporation of SRI can lead to high positive abnormal stock returns as simply as by following a long-short strategy, positive screening approach, or best-in-class approach. According to their research the best-in-class approach can lead to abnormal returns as high as +8.7% per year. However, when examining more of the academic literature, the majority of the studies argue that the incorporation of ESG criteria neither adds or destroys value. Thus, incorporating SRI does not lead to higher portfolio performance as so-called traditional investing. In addition, several of these earlier studies are subject to major limitations such as disparate methods and data.

More recent papers, for instance from Bebchuk, Cohen, & Wang (2013) as well as from Borghers, Derwall, Koedijk, & Ter Horst (2013) take these limitations into account and re-examine the performance of SRI under the basis of market learning as well as errors in expectations. Moreover, the paper of Borghers et al. (2013) find that even though SRI yielded abnormal stock returns in the past, it does not do so anymore after the year 2004. The reason for this phenomenon is mentioned to be the learning effect of market participants, meaning that the participants learn to price the information regarding SRI and ESG scores correctly. However, the most recent studies, for example from Lins et al. (2017) and Auer (2016) report new differing results conflicting with those of previously mentioned. Using new methods and data sets, Lins et al. (2017) find out that firms that entered the financial crisis period with high CSR ratings have significantly higher crisis-period stock returns compared to those who entered with low CSR ratings.

Next three subchapters will compare in more detail whether SRI outperforms, underperforms, or does neither compared to a risk-free investment by going through some

of the relevant studies on the topic. The subchapters are divided into categories of positive effect, negative effect, and insignificant effect, respectively.

Positive effect

Nofsinger and Varma (2014) find in their study that socially responsible mutual funds outperform conventional mutual funds during the times of market crises. They are capable of providing evidence through CAPM, Fama and French three-factor model as well as through Carhart (1997) four-factor model. Their study shows that during times of macroeconomic crises, socially responsible mutual funds are able to outperform conventional mutual funds by 1.18% annually, as for during “ordinary” market times, these conventional mutual funds outperform. Furthermore, the authors find that socially responsible funds outperform between 2000 and 2011, however with the result that they underperform conventional funds during ordinary times. They argue that mutual funds, which are especially focused on the ESG dimensions and use positive screening as a strategy, are driving an “asymmetric return pattern”. More specifically, they examine whether socially responsible funds are able to limit the investment’s downside risk by applying different ESG screening approaches, that is, during ordinary market times as well as during times market crises.

Furthermore, the authors examine the reasons behind the fact that the SRI assets under professional management increased by 380% between 1995 and 2010, meaning a significant annualized growth rate of 25.33%, even though previous academic research is indicating SRI to be unfavorable and expensive. Therefore, the authors state that there has to have a some sort of utility that the socially responsible investors are deriving from their investment approach. However, it shown that not every SRI approach has the same impact, but the focus needs to be put for the screening approach of ESG dimensions. For instance, the authors find that companies with good levels of corporate governance seem to overperform the other screening approaches during periods of market crises. Finally, the authors find that positive ESG screening approaches seems to lead to better as well as more sustainable returns compared to negative screening approaches.

Negative effect

In their paper, Halbritter and Dorfleitner (2015) investigate the link between the social and financial performance of corporations based on their ESG ratings gathered from data sources such as KLD, Bloomberg, and ASSET4 (which is also used in this thesis). The study focuses on the U.S. markets between 1991 and 2012, and applies Carhart (1997) four-factor model and cross-sectional Fama and MacBeth (1973) regressions in its econometric framework through ESG portfolio approach. Where the previous empirical research shows a relationship between positive abnormal returns and ESG ratings, Halbritter et al. demonstrate that the ESG portfolios, in fact, do not yield any abnormal returns when comparing companies with high and low ESG ratings. Even though the regression analysis derived from Fama and MacBeth (1973) disclose a considerable impact of various ESG factors, investors are barely able to utilize this relationship. Furthermore, in their study, Halbritter et al. employ the same approach as Kempf and Osthoff (2007) in theirs, grouping the companies into portfolios by their ESG scores. That is, taking the top 20% and bottom 20% of the companies ranked by their ESG scores, then buying the best ranked and short selling the worst ranked, i.e. using the long-short approach. All in all, the results of the study indicate that investors should not anymore anticipate abnormal stock returns by trading with the long-short approach of ESG scores.

Insignificant effect

As many other studies, Hamilton, Jo, and Statman (1993) argue that some investors favor certain types of firms over others according to criteria such as weapon production. In other words, these investors seem to be avoiding sin stocks, and are thus categorized as socially responsible investors. Hamilton et al. find that SRI mutual funds do not earn any statistically significant abnormal risk-adjusted returns, and that performance of these SRI funds does not differ statistically from the performance of conventional mutual funds. The authors present three different hypotheses that create the framework for future studies about SRI: 1) the risk-adjusted expected returns of SRI and conventional portfolios are *equal*; 2) the risk-adjusted expected returns of SRI portfolios are *lower* than those of

conventional portfolios; and 3) the risk-adjusted expected returns of SRI portfolios are *higher* than those of conventional portfolios.

Furthermore, a somewhat more recent study of Bauer, Koedijk, and Otten (2005) find that there is no significant differences between the risk-adjusted returns of SRI, or ethical, funds and conventional funds by examining a sample of 103 US, UK, and German ethical mutual funds during a period between 1990 and 2001. As many other previously mentioned studies, they use the Carhart (1997) four-factor model, and are hereby one of the first to overcome the benchmark problem that most prior SRI studies suffered from that time. The results of the study also indicate, that during the sample period, SRI mutual funds experienced a catching up phase before yielding returns similar to the conventional funds.

Lastly, one of the more recent studies of Belghitar, Clark, and Deshmukh (2014) also suggest that there is no significant difference between the performance of socially responsible investments and conventional investments. By using previous research and empirical mean-variance evidence, the authors find the results to be truly insignificant. Belghitar et al. find significant evidence that socially responsible investing is, in fact, rather costly as these socially responsible investors lose the additional expected yield in transaction fees.

3. DATA AND METHODOLOGY

This chapter explains the sources of data as well as the methodology used in this thesis. The first section aims to describe the overall data and all the available data sources used. Furthermore, the next section describes the portfolio construction process in detail, following the methodology used as well as the description of performance measurement and descriptive statistics. Lastly, the methodology is described in more detail by providing theories on Capital Asset Pricing Model, Fama and French (1993) three-factor model, and Fama and French (2015) five-factor model, as all of them are used to execute the regression analysis of this thesis.

3.1. Data sources and description

There are various possible data sources in academic literature when working with ESG data. When observing academic literature on ESG, it can be noted that the most used data sources used are mostly likely the “MSCI ESG Ratings” (MSCI) and the database of “Kinder, Lydenberg, Domini & Co.” (KLD). Other less frequently used and mentioned databases in academic literature are for instance the “Thomson Reuters ASSET4” (henceforth “ASSET4”) as well as the well-known and reputable Bloomberg. Due to the restrictions in data accessibility, and because ASSET4 ESG data is only rarely cited among academics compared to the dominant MSCI and KLD data, this thesis will employ the Thomson Reuters ASSET4 ESG database as it is also providing the equal level of suitability.

According to “Thomson Reuters ESG Scores” (2019) by Thomson Reuters, ASSET4 ESG scores are intended to measure a company’s relative performance by ESG scores compared to its peers. In order to do so, Thomson Reuters uses publicly available data such as media & non-governmental organization (NGO) reports as well as company disclosures, to capture over 400 company level ESG metrics in order to conduct a detailed assessment of each company. The 178 most relevant data points are chosen for the scoring process by comparability, data availability, and industry relevance. These data points are

then grouped into ten categories, each related to one of the three fields: “Environmental”, “Social”, or “Governance”. The final ESG scores are calculated by percentile rank scoring, considering the amount of companies that are worse and equal to the analyzed one, but also all of the companies that have no scoring at all. These annually updated ESG scores by a fiscal year, can rank from 0 to 100, and are meant to be interpreted as percentile values, thus reflecting how well (high score) or how bad (low score) a company’s ESG performance is relative to its peers. (Thomson Reuters 2019.)

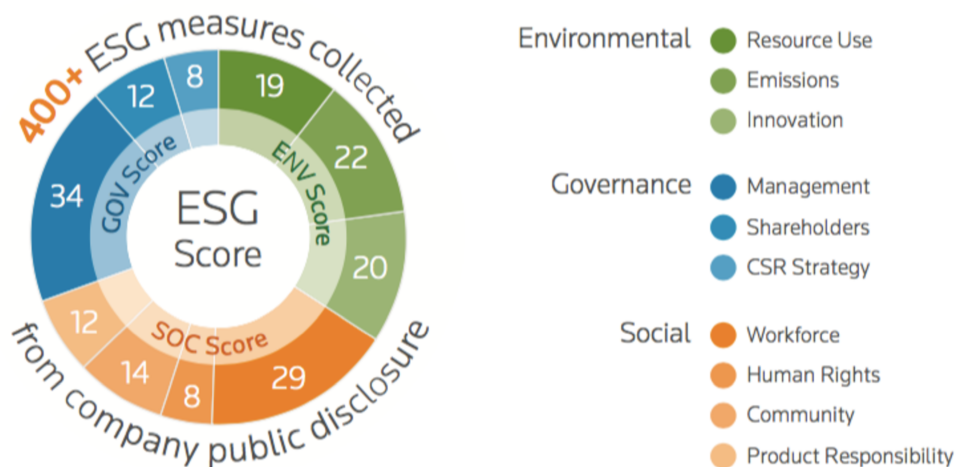


Figure 6.) The Thomson Reuters ASSET4 ESG measures and score formation (Thomson Reuters 2019).

Continuing, the data collected from the Thomson Reuters’ database is divided into two parts. The first part of the retrieved data consists of the yearly, end-of-the-year, combined ESG scores as well as the individual “Environmental”, “Social”, and “Governance” scores of all the selected financial companies listed in the NYSE between January 2002 and January 2017. Years 2018 and 2019 are not included in the data sample of this thesis, as the ESG data was just partially available or not available at all. The second part of the data comprises of the monthly, end-of-the-month, closing prices of all the same companies mentioned previously. This results in approximately 35,000 monthly observations, covering exactly 181 months over a 16 year period, thus spanning 193 financial companies that are part of the NYSE investment pool at the given sample period. Furthermore, these monthly closing price observations are condensed into yearly holding period returns in order to make the regression analyses more straightforward. Moreover,

in the data set collected, the financial companies are divided into two industry categories: banks and financial services companies. This means that for instance insurance companies, private equity firms, REITs, etc. are excluded from this specific set of data.

To summarize, the majority of the data is collected from the ASSET4 database. That is, the monthly closing prices as well as the annual combined and individual ESG scores are all retrieved from this database. However, all the necessary data needed for the regression analyses, i.e. for the implementation of CAPM, three-factor model, and five-factor model are all gathered from the Kenneth R. French's (2019) database. These are the size factor (SMB_t), value factor (HML_t), profitability factor (RMW_t), and investment factor (CMA_t) data, as well as the risk-free rates of return.

The following figure shows the distribution of combined ESG scores as well as the distribution of the individual “Environmental”, “Social”, and “Governance” scores of the selected financial companies over the whole sample period. The vertical axes of each histogram indicate the number of companies that experience a specific score interval of ten points, whereas the horizontal axes represent that specific score interval.

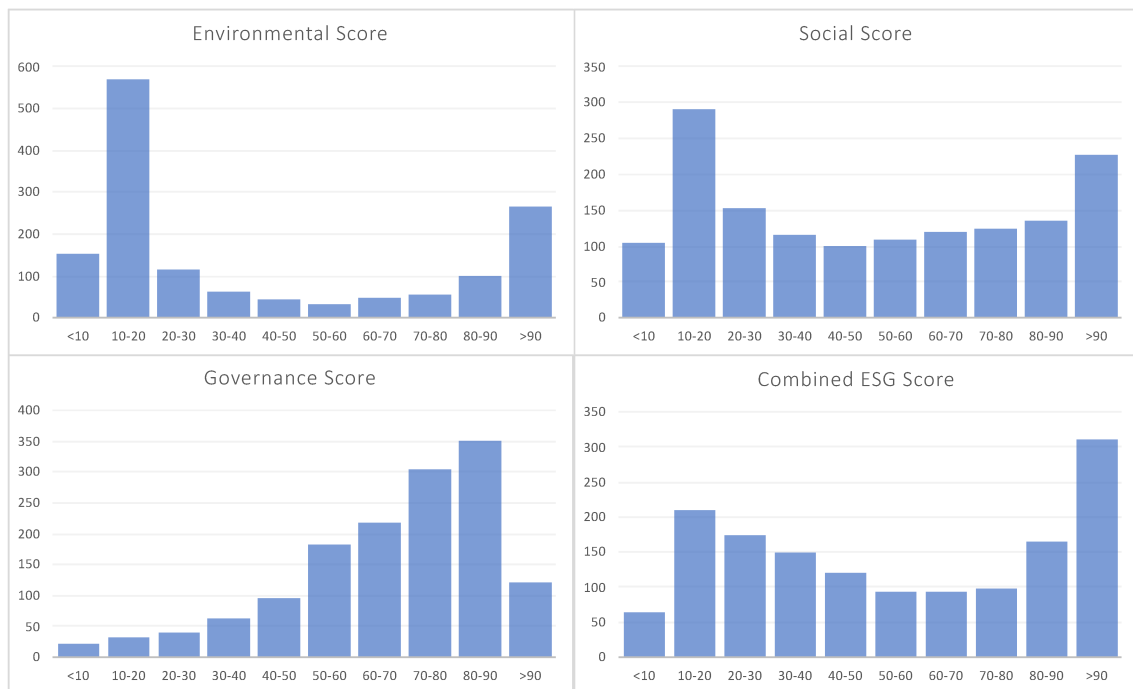


Figure 7.) ESG score distribution of the selected companies over the whole sample period between 2002 and 2017.

As can be seen from the histogram, the distribution of “Environmental” scores shows that most of the selected financial companies exhibit very low, or alternatively very high, environmental scores. To be more specific, the majority of the companies exhibit very low environmental scores between the 10th and 20th percentiles as well as very high scores around the 90th percentile. A possible implication for this phenomenon is that these companies either tend to fully invest their resources into environmental dimensions, or vice versa, tend to exclude environmental dimension totally from their corporate strategy. This is not rather surprising, firstly, because the financial industry is commonly considered as rather environmentally unconscious, and secondly, since the sample period covers the time before the financial crisis as well as the time after the crisis. Therefore, one could assume that the financial sector has become more environmentally conscious over the past decade, thus indicating that many of these companies rank around the 90th percentile in addition to the majority that counts among the 10th and 20th percentiles. All in all, the histogram could be a clear indication that the financial sector has become more aware over the past decade when considering environmental dimensions, yet it is impossible to tell as the figure does not take time periods into account.

Continuing, the histogram of “Social” score distribution illustrates that the social scores are rather evenly distributed, however also with the tendency to have scores around the 10th and 20th percentiles as well as 90th percentile, yet not to the same extent as in the environmental dimension. Since the scores are rather evenly distributed among the social dimension, it is inconvenient to draw any conclusions, as the variation in this dimension is high. Furthermore, when observing the distribution of “Governance” scores, it can be noticed that the majority of the financial companies obligate themselves to maintain good standards in corporate governance. This can be illustrated by the fact that the histogram is ascending towards the best-in-class percentile and therefore most of the firms are placed between the 50th and 90th percentile. However, this is not rather surprising since good corporate governance as a dimension of corporate responsibility has been around the longest when comparing to the social and environmental dimensions. Thus, it can be stated that a high level of corporate governance is a significant factor of corporate responsibility among the financial sector. Finally, the combined ESG score histogram, which represents the equally weighted average across the three individual dimensions,

illustrates that the financial companies tend to have either very high ESG scores of over the 90th percentile or alternatively quite low scores around the 10th and 40th percentile. Hereby, the majority of the financial companies can be categorized either to the best-in-class or to the worst-in-class groups.

Supporting the findings above, table 1. contains the descriptive statistics of individual “Environmental”, “Social”, and “Governance” scores as well combined ESG scores. The statistics cover the whole sample period from 2002 to 2017, hereby including 5,826 individual or combined end-of-the-year ESG observations.

	Mean	Median	Max	Min	STD	Skewness	Kurtosis
Env. Scores	40.750	21.582	95.262	11.675	33.310	0.680	-1.215
Soc. Scores	51.024	49.803	97.345	6.965	30.538	0.072	-1.404
Gov. Scores	68.616	72.577	96.089	13.443	19.035	-0.955	0.820
ESG Scores	55.712	53.658	96.977	6.990	30.089	0.003	-1.481

Table 1.) Descriptive statistics of the ESG scores over the whole sample period between 2002 and 2017.

As can be noted from the table, the average “Governance” scores for the financial companies listed in the NYSE between 2002 and 2017 tend to be approximately 17.6 percentage points higher than the “Social” scores and approximately 27.9 percentage points higher than the “Environmental” scores. Supporting the findings from the previous histograms, the distribution of “Governance” scores is upward sloping with a median of 72.6, whereas the distribution of “Environmental” scores tend to focus on the low ends with a median of 21.6. The “Social” scores are more evenly distributed with a median of 49.8 and a mean of 51.0. Furthermore, the standard deviation (“STD”) is the lowest for “Governance” scores, as it tends to be approximately 10 to 14 percentage points lower compared to the other ESG dimensions.

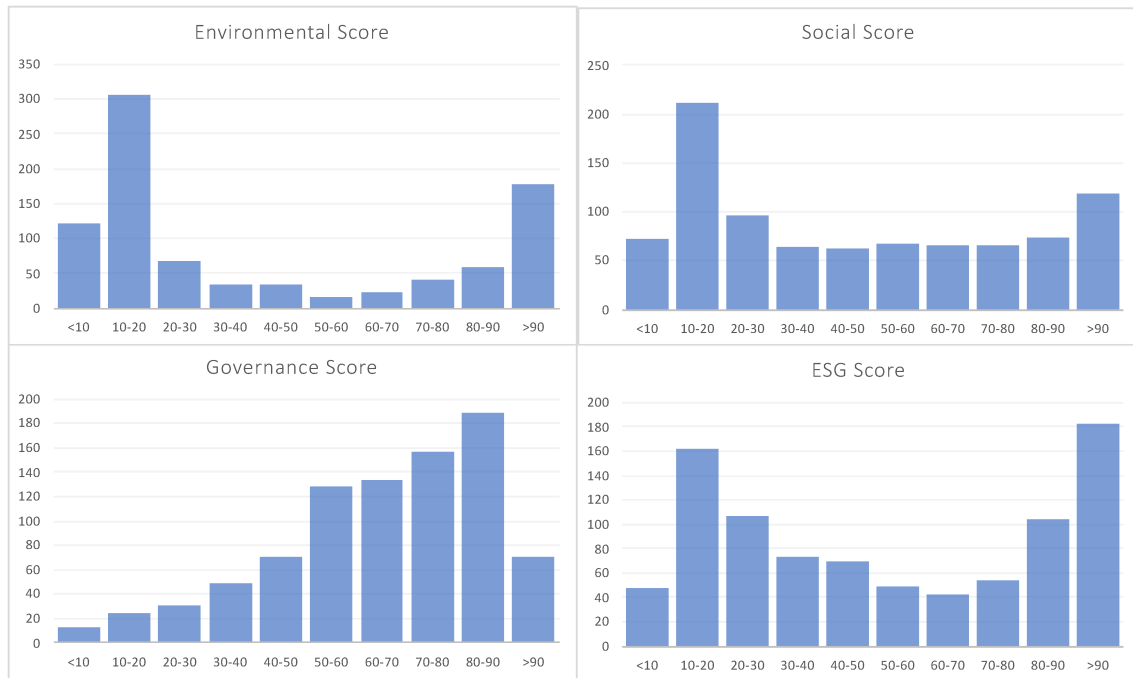


Figure 8.) ESG score distribution of the selected companies over the post-crisis sample period between 2010 and 2017, i.e. after the financial crisis.

Continuing, figure 8. shows the ESG score distribution of the selected companies over the post-crisis sample period after the financial crisis, i.e. between 2010 and 2017. As before, the vertical axes of each histogram indicate the number of companies that experience a specific score interval of ten points, whereas the horizontal axes indicate that specific score interval ranging from 0 to 100.

As figure 8. illustrates, the distribution of “Environmental” scores indicates that most of the selected financial companies exhibit either very low or very high environmental scores during the 8-year time period. As mentioned previously, a possible implication for this phenomenon is that these companies either tend to fully invest their resources into environmental dimensions, vice versa. This is not surprising, as the financial industry is commonly considered as rather environmentally unconscious. Furthermore, as argued earlier over the whole sample period of 16 years, one could assume that the financial sector has become more environmentally conscious over the past decade. However, it can be noted that the financial sector has definitely not become more aware over the past decade when considering environmental dimensions, as it can clearly be seen from the histogram that this is not the case with this sample.

Consistent with the whole sample, the post-crisis period's histogram of "Social" score distribution also demonstrates that the social scores are rather evenly distributed. However also with the tendency to have scores around the 10th and 20th percentiles as well as 90th percentile, yet not to the same extent as in the environmental dimension. Overall, a simple conclusion can be drawn that the selected financial companies are not interested in incorporating social dimensions, as it can be clearly seen from the high tendency to have social scores around the 10th and 20th percentiles. Moreover, when examining the "Governance" score distribution, it can be noted that most of the financial companies obligate themselves to maintain good or great standards in corporate governance. This can be demonstrated by the fact that the histogram is ascending towards the best-in-class percentile and therefore most of the firms are placed between the 50th and 90th percentile. However, as already mentioned, this is not surprising since corporate governance dimension has been around the longest when comparing to the environmental and social dimensions. Lastly, the combined ESG score histogram demonstrates that the financial companies tend to have either very high ESG scores of around the 80th and 90th percentile or alternatively rather low scores around the 10th and 30th percentile. Therefore, also in the post-crisis period the majority of the financial companies are categorized either to the best-in-class or to the worst-in-class groupings.

Supporting the observations argued above, table 2. comprises the descriptive statistics of individual as well combined ESG scores. The statistics cover the post-crisis sample period of eight years, i.e. a period after the financial crisis between 2010 and 2017. Hereby, containing a total of 3,532 individual or combined end-of-the-year ESG observations.

	Mean	Median	Max	Min	STD	Skewness	Kurtosis
Env. Scores	42.132	22.174	94.664	9.513	35.125	0.512	-1.521
Soc. Scores	65.991	68.508	96.601	11.826	19.950	-0.691	0.143
Gov. Scores	46.602	41.102	96.256	6.973	30.596	0.264	-1.410
ESG Scores	53.120	49.814	96.275	5.996	31.419	0.071	-1.584

Table 2.) Descriptive statistics of the ESG scores over the post-crisis sample period between 2010 and 2017.

As table 2. illustrates, the average “Social” scores for the selected financial companies between 2010 and 2017 tend to be approximately 24 percentage points higher than the “Environmental” scores and around 19 percentage points higher than the “Governance” scores. This is quite an opposite compared to the ESG score distribution over the whole sample period of 16 years, where the average “Governance” scores tend to be significantly higher than the “Social” and “Environmental” scores. Furthermore, supporting the findings from the previous histograms, the distribution of “Social” scores tend to be negatively skewed with a skewness of -0.691. Also the distribution of “Environmental” scores tend to focus on the low ends with a median of 22.2 and a mean of 42.1. Furthermore, the standard deviation is the lowest for “Social” scores, as it tends to be around 10 to 15 percentage points lower compared to the other ESG dimensions.

3.2. Portfolio construction and descriptive statistics

In the portfolio construction process, the portfolios are constructed using a screening approach at the beginning of each year. This means that the portfolios are created by using either a best-in-class (positive screening) or a worst-in-class approach by screening 20% of the best and 20% of the worst ranked financial stocks grouped by their ESG scores. More accurately, the firms are grouped at the beginning of each year by their combined ESG scores as well as individual “Environmental”, “Social”, and “Governance” scores. Thus, eight different types of portfolios are constructed: two portfolios are created for the combined ESG scores (20% of the best and worst) and two portfolios (20% of the best and worst) for each of the individual “Environmental”, “Social”, and “Governance” dimensions. As already mentioned, the portfolios are created by using the information available at the beginning of each sample period. They are then held until the end of the observed year, until new information about the ESG scores are announced and the portfolios are then re-balanced appropriately.

To clarify the portfolio construction process even more, this paragraph presents the in-depth portfolio creation process for the whole sample period and in a form of an example. Firstly, by using the ESG scores from the beginning of the year of 2002, all of the selected

financial companies listed in the NYSE at that specific point in time are grouped from the highest to the lowest according to their respective “Environmental”, “Social”, “Governance”, and “ESG” scores. Secondly, 20% of the firms that are exhibiting the highest individual and combined scores respectively, are then grouped into separate portfolios. Similarly, 20% of the firms that are exhibiting the lowest scores are also grouped into separate portfolios. As mentioned in the previous paragraph, these portfolios are then held until the end of the observed year, until new information about the ESG scores are announced and the portfolios are then re-balanced appropriately. This same process is duplicated until the end of the whole sample period in 2017, thus resulting in eight separate portfolios: top & bottom “Environmental”, top & bottom “Social”, top & bottom “Governance”, and top & bottom “ESG”. Furthermore, from now on the portfolios comprising 20% of the highest (lowest) scored firms in any of the four dimensions are called “Top” (“Bottom”). For instance, “Soc. Top” refers to a portfolio containing 20% of the firms with the highest “Social” score rating, and “Gov. Bot” refers to a portfolio containing 20% of the firms with the lowest “Governance” score rating.

Table 3. provides the descriptive statistics of the annual excess returns for the whole sample period. The excess returns are calculated by subtracting a risk-free rate of return from a holding period return of one year, that is $R_{it} - R_{Ft}$. As mentioned earlier, the annual risk-free rates are gathered from the Kenneth R. French’s (2019) web page. Furthermore, when calculating the holding period returns (HPRs) that are comprised from the monthly closing prices, cash dividends or dividend yields are not taken into account, whereupon the formula can be put as follows (Bodie et al. 2014: 128):

$$(7) \quad HPR = \frac{P_1 - P_0}{P_1},$$

where: P_1 = Ending price of a share
 P_0 = Beginning price of a share

Table 3. presents the descriptive statistics of the annual excess returns for the whole sample period, hereby covering 181 months of monthly return observations, condensed into 16 years of holding period returns, spanning from 2002 to 2017. “Env.” expresses

that the portfolios are created by using Environmental scores as the determining criteria, whereas “Soc.” and “Gov.” indicate that the portfolios are constructed by employing Social and Governance scores as the determiners. Moreover, “ESG” naturally expresses that the portfolios are created by using the combined ESG scores as the determining criteria. Lastly, “Top” (“Bottom”) indicates that the portfolios are created by using the best-in-class (worst-in-class) approach, i.e. screening 20% of the best (worst) performing selected financial companies listed in the NYSE by their individual Environmental, Social, and Governance scores as well as combined equally weighted ESG scores.

	Mean	Median	Max	Min	STD	Skewness	Kurtosis
Env. Top	5.970	9.289	78.836	-78.808	30.826	-0.271	0.288
Env. Bottom	10.238	6.952	161.051	-82.994	35.227	0.516	1.775
Soc. Top	9.157	10.510	78.836	-70.157	28.859	-0.058	0.703
Soc. Bottom	9.681	6.555	192.036	-86.085	43.672	0.902	2.743
Gov. Top	7.176	8.894	118.337	-78.808	31.190	-0.107	1.279
Gov. Bottom	7.531	7.071	106.128	-68.795	34.046	0.254	0.045
ESG Top	9.299	10.242	78.836	-78.808	29.080	-0.277	0.817
ESG Bottom	8.302	9.602	196.836	-84.485	44.927	0.749	2.382

Table 3.) Descriptive statistics of the annual excess returns over the whole sample period between 2002 and 2017.

Table 3. indicates mixed results of the annual excess returns when examining under the first alternative hypothesis (H_1) that incorporating high ESG criteria leads to positive abnormal stock returns in the financial sector. Firstly, when observing the combined equally weighted ESG scores, it can be noticed that the high ESG scored financial companies, in fact, overperform the low scored ones. The mean annual excess returns for the best-in-class ESG portfolio is around 9.3%, whereas the mean annual excess returns for the worst-in-class ESG portfolio is 8.3%. In other words, the 20% of the best ranked financial stocks grouped by their combined ESG scores seem to overperform the worst ranked ones by around 1.0% annually. However, when examining more closely, it can be noted that the excess annual returns categorized by their individual ESG dimensions seem to give opposite results: the worst-in-class portfolios seem to overperform the best-

in-class portfolios. The mean annual excess returns for the best-in-class Environmental portfolio is 5.97%, whereas the mean excess returns for the worst performing financial firms ranked by their Environmental score is around 10.24%. This indicates that the worst-in-class Environmental portfolio would overperform the best-in-class one by significant 4.27% annually. Furthermore, the worst-in-class Social and Governance portfolios seem to overperform the best-in-class portfolios as well. The mean annual excess returns for the best-in-class Social portfolio is approximately 9.16%, whereas the mean excess returns for the worst-in-class Social portfolio is around 9.68%. That is, the bottom 20% of the financial companies ranked by their individual Social scores seem to overperform the top 20% performing ones by 0.52% annually. Lastly, the mean excess returns for the best-in-class Governance portfolio is around 7.18% and the annual mean excess returns for the worst-in-class Governance portfolio is around 7.53%, hereby stating that bottom 20% performing financial firms categorized by their Governance scores seem to overperform the best-in-class Governance portfolio as well.

The median excess annual returns of the selected financial companies are ranging between 6.56% and 10.51%. “Social Top” portfolio expresses the highest whereas “Social Bottom” portfolio expresses the lowest median excess returns of the sample. Furthermore, the maximum annual excess returns are ranging between 78.84% and 196.84%, and the minimum excess returns are varying from -86.09% to -68.80%. Altogether, the standard deviation of returns (volatility) seems to be ranging from 28.86% to 44.93%, thus varying around 35% on average. Lastly, the descriptive statistics’ table shows that the excess return distributions of the “Social Bottom” and the “ESG Bottom” portfolios are the most positively skewed, whereas the return distributions of the “Environmental Top” as well as the “ESG Top” portfolios are the most negatively skewed. Moreover, the “Social Bottom” portfolio seems to have the highest and “Governance Bottom” the lowest kurtosis of the sample.

Continuing, table 4. presents the descriptive statistics of the annual excess returns for the post-crisis sample period, i.e. period after the financial crisis. Hereby, covering 96 months of monthly return observations, condensed into 8 years of HPRs, spanning from 2010 to 2017. As earlier, “Env.”, “Soc.”, and “Gov.” expresses that the portfolios are created by

using Environmental, Social, and Governance scores as the determining criteria, whereas “ESG” expresses that the portfolios are created by using the combined ESG scores as the determining criteria. “Top” (“Bottom”) indicates that the portfolios are created using the best-in-class (worst-in-class) approach, i.e. screening 20% of the best (worst) performing financial companies by their individual as well as combined ESG scores.

	Mean	Median	Max	Min	STD	Skewness	Kurtosis
Env. Top	8.527	10.327	62.626	-46.264	22.419	-0.079	0.112
Env. Bottom	16.847	12.870	161.051	-71.044	33.837	0.939	2.868
Soc. Top	11.319	12.928	70.038	-38.960	22.615	0.186	0.412
Soc. Bottom	14.316	10.727	161.051	-71.044	39.615	0.824	2.021
Gov. Top	11.940	13.232	62.626	-53.861	23.322	-0.412	0.583
Gov. Bottom	9.513	11.333	106.128	-60.982	33.271	0.234	0.178
ESG Top	11.526	12.666	62.626	-44.416	21.194	-0.158	0.295
ESG Bottom	11.119	13.536	161.111	-72.768	40.633	0.547	1.767

Table 4.) Descriptive statistics of the annual excess returns over the post-crisis sample period between 2010 and 2017.

Also table 4. indicates mixed results of the annual excess returns when examining under the thesis’ first alternative hypothesis (H_1) that incorporating high ESG criteria leads to positive abnormal stock returns in the financial sector. When examining the combined equally weighted ESG scores, it can be noticed that the high ESG scored financial companies, in fact, overperform the low scored ones. The mean annual excess returns for the best-in-class combined ESG portfolio is around 11.5%, whereas the mean annual excess returns for the worst-in-class combined ESG portfolio is 11.1%. In other words, the 20% of the best ranked financial stocks grouped by their combined ESG scores seem to overperform the worst ranked ones by around 0.4% annually.

However, when observing more closely, it can be noticed that the excess annual returns categorized by their individual ESG dimensions seem to give rather opposite results. For instance, the mean annual excess returns for the best-in-class Environmental portfolio is 8.53%, whereas the mean excess returns for the worst-in-class portfolio is 16.85%. This

indicates that the worst-in-class Environmental portfolio would overperform the best-in-class portfolio by significant 8.32% annually. Also, the worst-in-class Social portfolio seems to overperform the best-in-class portfolio as well. The mean annual excess returns for the best-in-class Social portfolio is 11.32%, whereas the mean excess returns for the worst-in-class Social portfolio is around 14.32%. In other words, the bottom 20% performing financial companies ranked by their individual Social scores seem to overperform the top 20% performing companies by 3.00% annually. Nevertheless, the annual excess returns for the Governance portfolios differ from the two previous. The mean annual excess returns for the best-in-class Governance portfolio is 11.94%, whereas the mean annual excess returns for the worst-in-class Governance portfolio is 9.51%, stating that the best-in-class portfolio overperforms by 2.43% annually. Overall, the mean annual excess return for all the created portfolios seem to be higher in the post-crisis sample period between 2010 and 2017. This is intuitively rather straightforward, as during the financial crisis the NYSE Composite Index crashed from its then all-time-highs at the end of 2007 for almost 55% till the February of 2009 (New York Stock Exchange 2019).

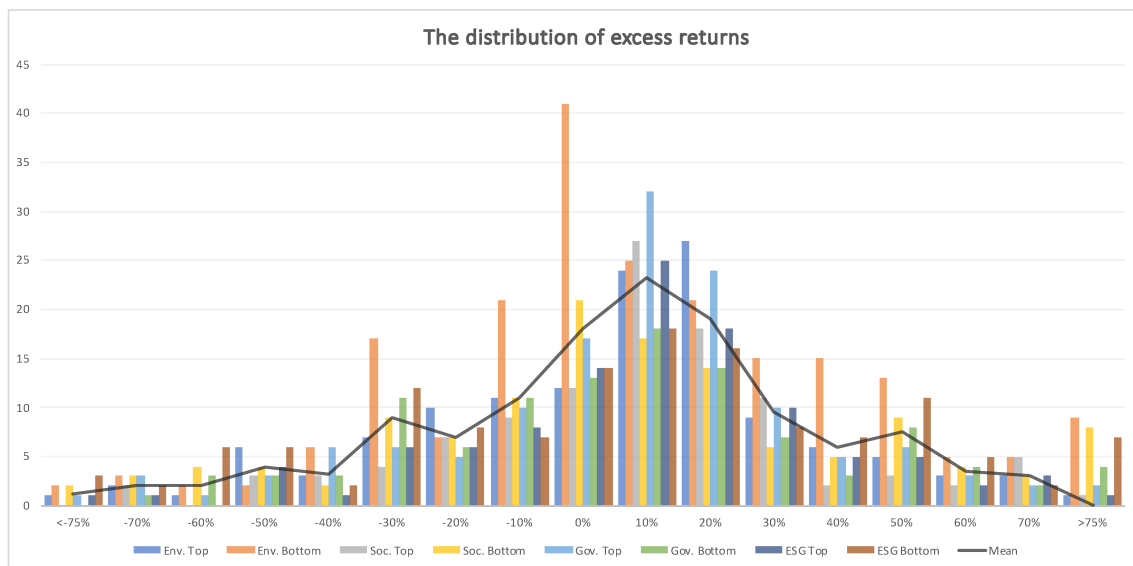


Figure 9.) The overall distribution of the annual excess returns for all the created best-in-class and worst-in-class portfolios sorted by their individual and combined ESG dimensions.

Continuing, figure 9. compiles the annual excess returns of all the created ESG portfolios over the whole sample period between 2002 and 2017. The vertical axis indicates the number of excess returns that belong to a specific return interval, whereas the horizontal

axis represents these intervals (10 pp) of excess returns. The black line illustrates the mean of the portfolios' excess returns, therefore indicating that the created portfolios' annual excess returns are somewhat normally distributed, if clearly skewed. Furthermore, there seems to be also rather clear deviations from the normal in the -30% and +50% areas of excess returns.

3.3. Methodology and performance measurement

This thesis' empirical part seeks to analyze whether the incorporation of ESG criteria has any statistically significant positive or negative impact on selected financial companies' stock returns. To be more specific, the goal is to analyze the performance of NYSE's financial sector by screening the stocks in this specific investment universe for their ESG scores. In order to measure this performance, this thesis will use the CAPM, Fama and French (1993) three-factor model as well as the Fama and French (2015) five-factor model in its regression analyses.

The regression analyses covers a data sample of banks and financial services companies listed in the NYSE, covering from January 2002 to January 2017, thus spanning a period of 16 years. This represents a sample of 181 months, covering 193 companies in the financial sector that are, or were, listed in the NYSE during the mentioned period. The empirical analyses is divided into two separate parts: to the analysis of the whole sample period (2002-2017) and post-crisis sample period (2010-2017). Moreover, the OLS regression analyses will determine if a best-in-class approach, which consists of yearly re-balanced portfolios of the 20% highest ESG scored financial stocks, yield any abnormal stock returns. Moreover, the analyses will also examine the 20% worst ranked financial companies (worst-in-class), and compare them to the top 20% best scored ones.

3.3.1. Capital Asset Pricing Model

As already mentioned previously in this thesis, the Capital Asset Pricing Model (CAPM) is an asset pricing model that demonstrates the relationship between systematic risk and

expected return for a certain asset or a portfolio. CAPM includes various components such as the risk-free rate of return, the systematic risk, and the rate of return for a market portfolio. As an outcome it aims at finding the expected rate of return for a specific asset or a portfolio. (Fama & French 2004; Sharpe 1964.) It is critical to understand the theories behind CAPM as the Fama and French (1993) three factor model as well as the Fama and French (2015) five-factor model, which are also used to execute the regression analyses of this thesis, are based on it.

Criticism towards CAPM

Academic literature presents a great amount of criticism towards the CAPM. According to Roll (1977) it is rather unthinkable to ever even observe a real market portfolio, as it does not exist, meaning that using CAPM in practice is simply impossible. Furthermore, Eugene Fama and Kenneth French presents even more serious criticism towards the model. According to Fama and French (1992) there is not any relationship between the beta coefficient (market risk) and the expected rate of return after certain company-specific factors are taken into account. Furthermore, they state that the beta coefficient alone is not capable to explain the expected returns, yet there are other factors effecting it as well. Thus, Fama and French (1993) introduces the well-known three-factor model, which adds size (SMB_t) and value (HML_t) factors to the market risk (β_i) factor of the CAPM. Later on the five-factor model (Fama & French 2015) is introduced, which again adds two new additional factors into the previous model: profitability and investment. Altogether, all of these three models will be used to execute the regression analyses of this thesis, yet the five-factor model is supposed to give the most relevant and CAPM the least relevant results.

3.3.2. Fama and French multi-factor models

The Fama and French three-factor model

Again, the Fama and French (1993) three-factor model is designed to describe stock returns, complementing the Capital Asset Pricing Model. The three-factor model was

established to take into account the factors that are not explained by the CAPM. These are, the excess return on a market portfolio ($R_{MT}-R_{FT}$), the return on a diversified portfolio of small stocks minus the return on a diversified portfolio of big stocks (SMB_t) as well as the difference between the returns on diversified portfolios of high and low book-to-market stocks (HML_t). (Fama et al. 1993; 2015.)

The Fama and French five-factor model

The five-factor model complements the previously mentioned three-factor model by adding two new additional factors in it; RMW_t , or “Robust Minus Weak” and CMA_t , or “Conservative Minus Aggressive”. The first one measures the difference between the returns of a diversified portfolio consisting of stocks with high and low profitability. The latter measures the difference between the returns of a diversified portfolio consisting of companies with low investment rate and high investment rate. According to Fama and French (2015), this five-factor asset pricing model is capable of explaining up to 94% of the cross-section variance of the observed portfolios’ returns, and thus is expected to provide the most relevant results within the all asset pricing models. (Fama et al. 2015.)

3.3.3. Performance measurement

In the thesis, $R_{it}-R_{Ft}$ measures the excess returns of the previously created ESG scored portfolios. It is calculated by using the risk-free rate of 1-month T-Bill, together with the factors of SMB_t , HML_t , RMW_t , and CMA_t gathered from Kenneth R. French’s (2019) web page. The CAPM, the three-factor model as well as the five-factor model all will be used in the OLS regression analyses of this thesis in order to determine whether the portfolios grouped on ESG scores deliver any statistically significant returns (α_i) that cannot be explained by the beta coefficient factors. That is, the market factor of the CAPM, the three factors in the three-factor model, and the five factors in the five-factor model. If the results denote that the incorporation of ESG criteria can be used to accomplish statistically significant abnormal stock returns, it would support the hypothesis (H_1) that incorporating high ESG criteria leads to positive abnormal stock returns in the financial sector, vice versa (H_2).

4. EMPIRICAL RESULTS AND ANALYSIS

This chapter presents the empirical results gathered from the OLS regression analyses. The OLS regression results are implemented by using three different regression models: the “one-factor” Capital Asset Pricing Model, the Fama and French (1993) three-factor model, and the Fama and French (2015) five-factor model, as described earlier in this thesis. Chapter 4.1. presents the empirical results over the whole sample period between 2002 and 2017, whereas chapter 4.2. presents the results on the post-crisis period between 2010 and 2017. Furthermore, this chapter examines the gathered results under the H_1 : “incorporating high ESG criteria leads to *positive* abnormal stock returns in the financial sector” as well as under the H_2 : “incorporating high ESG criteria leads to *negative* abnormal stock returns in the financial sector”. Therefore, aiming to reject the null hypothesis H_0 : “incorporating high ESG criteria leads to neither positive nor negative abnormal stock returns in the financial sector”. In conclusion, chapter 4.3. then summarizes all the gathered empirical results obtained from the OLS regression analyses as well as provides the final analysis of the results.

Continuing with the terminology and abbreviations, “Alpha” (α) indicates the estimated coefficient, that is, the abnormal returns that cannot be explained by the beta coefficient factors in the CAPM, three-factor model and/or in the five-factor model. Thus, possibly offering information whether the implementation of ESG criteria has any effect on financial companies stock performance or not. The abbreviations $R_m - R_f$, SMB, HML, RMW, and CMA signify the Fama and French’s five different factors of beta coefficients, as described earlier in the thesis. R-squared (“R²”) measures the proportion of the variance for the dependent variable, i.e. $R_{it} - R_{Ft}$, that is explained by the independent variable(s), i.e. the beta coefficients, in the regression model. Hereby, indicating that a higher R-squared denotes a better model.

Furthermore, as previously stated, the data sample covers a total of 181 monthly return observations, thus resulting in 35,000 monthly observations, however condensed into 16 years of holding period returns, spanning from 2002 to 2017. “Env.” expresses that the portfolios are created by using Environmental scores as the determining criteria, whereas

“Soc.” and “Gov.” indicate that the portfolios are constructed by employing Social and Governance scores as the determiners. Moreover, “ESG” naturally expresses that the portfolios are created by using the combined ESG scores as the determining criteria. Lastly, “Top” (“Bottom”) indicates that the portfolios are created using the best-in-class (worst-in-class) approach. In other words, screening 20% of the best (worst) performing financial companies listed in the NYSE by their individual Environmental, Social, and Governance scores as well as combined equally weighted ESG scores. This description applies to all of the following OLS regression results presented.

Overall, the methodology and empirical framework of this thesis heavily complies with Derwall, Guenster, Bauer, and Koedijk (2005), Kempf and Osthoff (2007), Renneboog, Ter Horst, and Zhang (2008), as well as with Halbritter and Dorfleitner (2015), among many others. They all make use of the CAPM, Fama and French (1993) three-factor model, or the Carhart (1997) four-factor model, which are all prior versions of the Fama and French (2015) five-factor model that is used in the empirical part of this thesis along with the CAPM and three-factor model. Moreover, consistent with this thesis, Kempf et al. (2007) as well as Halbritter et al. (2015) use the best-in-class (worst-in-class) approach in their studies by screening for example 10%, 20%, and 25% of the best (worst) performing stocks sorted by their individual and combined ESG criteria.

4.1. Whole sample period

		Alpha	Rm-Rf	SMB	HML	R2
Env. Top	(1)	-5.940*** (0.01)	1.222*** (0.00)	0.388 (0.14)	0.542** (0.02)	0.940
	(2)	-4.727 (0.13)	1.272*** (0.00)			0.849
Env. Bottom	(1)	-2.006 (0.53)	1.067*** (0.00)	0.343 (0.39)	0.707** (0.04)	0.840
	(2)	-0.685 (0.86)	1.107*** (0.00)			0.703
Soc. Top	(1)	-3.503 (0.21)	1.217*** (0.00)	0.809** (0.03)	0.033 (0.90)	0.906
	(2)	-2.175 (0.50)	1.338*** (0.00)			0.842
Soc. Bottom	(1)	-2.456 (0.47)	1.220*** (0.00)	0.147 (0.72)	0.979*** (0.01)	0.855
	(2)	-1.150 (0.80)	1.227*** (0.00)			0.696
Gov. Top	(1)	-4.658* (0.08)	1.211*** (0.00)	0.618* (0.06)	0.166 (0.49)	0.916
	(2)	-3.489 (0.24)	1.301*** (0.00)			0.861
Gov. Bottom	(1)	-7.637** (0.02)	1.012*** (0.00)	0.921** (0.02)	0.287 (0.30)	0.880
	(2)	-5.854 (0.14)	1.145*** (0.00)			0.736
ESG Top	(1)	-3.051 (0.29)	1.142*** (0.00)	0.556 (0.12)	0.445 (0.12)	0.887
	(2)	-1.677 (0.63)	1.218*** (0.00)			0.790
ESG Bottom	(1)	-4.291 (0.34)	1.144*** (0.00)	0.419 (0.44)	0.931** (0.04)	0.778
	(2)	-2.602 (0.62)	1.192*** (0.00)			0.606

Table 5.) The OLS regression results over the whole sample period between 2002 and 2017, implemented with (1) the Fama and French (1993) three-factor model as well as with (2) the CAPM. Alpha (α) signifies the estimated coefficient, that is, the abnormal returns that cannot be explained by the beta coefficient factors. Thus, Rm-Rf, SMB, and HML signifies the factor loadings of the beta coefficients. R2 indicates the goodness-of-fit. The p-values are placed below the results, inside the parenthesis. *, **, and *** indicate the statistical significance at the 10%, 5%, and 1% level, respectively.

When investigating the previously presented OLS regression results, implemented with the CAPM and Fama and French (1993) three-factor model, under the first alternative hypothesis H_1 : “incorporating high ESG criteria leads to *positive* abnormal stock returns in the financial sector”, there are several findings that seem to disprove this hypothesis. First of all, all the portfolios grouped on the highest individual or combined ESG criteria, i.e. the best-in-class portfolios, indicate that the respective dimensions are negatively related to the stock returns. Moreover, of these portfolios “Environmental Top” is statistically significant at the 1% level and “Governance Top” at the 10% level. Less surprisingly, the portfolios sorted on the lowest ESG criteria also indicate that the respective dimensions are negatively related to the stock returns, of which the “Governance Bottom” portfolio is statistically significant at the 5% level. In conclusion, based on these findings, the first alternative hypothesis, H_1 , can be *rejected*, as it is obvious that incorporating high ESG criteria does *not* lead to *positive* abnormal stock returns in the financial sector when implemented with these particular methods and data.

Continuing with the analysis of the three beta coefficient factors used in the two regression models, i.e. $R_m - R_f$, SMB_t , and HML_t , some observations can be made. First of all, the market factor, $R_m - R_f$, is in all cases positively related to the stock performance and highly statistically significant at the 1% significance level, whether the OLS regression model is implemented with CAPM or with the Fama and French three-factor model. This demonstrates that the created portfolios’ excess returns, $R_{it} - R_{Ft}$, are in fact mainly driven by the markets. In addition, when observing the size (SMB_t) and value (HML_t) factors, some consistencies can be noticed. It seems that the excess returns of some portfolios are driven by the size factor as well as by the value factor. For example, it can be noted that the excess returns of the “Social Top” as well as “Governance Top” portfolios are positively related with the size factor, at the 5% and 10% levels of statistical significance, respectively. This observation demonstrates that small financial companies tend to have higher stock returns compared to large ones among these specific best-in-class portfolios. Furthermore, as mentioned, it also seems that the excess returns of some worst-in-class portfolios are driven by the value factor. It can be noticed that the excess returns of all the worst-in-class portfolios, excluding Governance, are positively related with the value factor, at the 1% or 5% levels of statistical significance. This indicates that

the worst-in-class portfolios' excess returns are partially driven by the value factor, demonstrating that value companies are outperforming growth companies in the financial sector. What comes to the R-squared ("R²") it can be distinctly stated the regression models' goodness-of-fit improves as more factors are incorporated into the model.

Furthermore, when investigating the OLS regression results under the second alternative hypothesis H_2 : "incorporating high ESG criteria leads to *negative* abnormal stock returns in the financial sector", there are findings that seem to accept the hypothesis. Firstly, all the portfolios grouped on the highest ESG criteria, i.e. the best-in-class portfolios, indicate that the respective dimensions are, in fact, negatively related to the stock returns. In addition, of these portfolios "Environmental Top" is statistically significant at the 1% level and "Governance Top" at the 10% level. Thus, 20% of the best financial companies ranked by their Environmental scores seem to generate annual abnormal stock returns of -5.94%, whereas the top financial companies ranked by their Governance scores seem to generate alpha of -4.66% after controlling the three beta coefficient factors. All in all, these findings indicate that in some cases the null hypothesis (H_0) gets rejected, and the second alternative hypothesis (H_2) hereby *holds*. Therefore, part of the results imply that incorporating high ESG criteria *does*, in fact, lead to *negative* abnormal stock returns in the financial sector when implemented with these particular methods and data.

Summarizing, it appears that these previous findings are in line with the paper of Renneboog et al. (2008), whom find that investors who are using ESG as an investment criteria tolerate a higher cost and are thus accepting inferior stock returns and overall financial performance (H_2 holds). Moreover, for example Halbritter et al. (2015) demonstrate that ESG portfolios do not yield any positive or negative abnormal returns when comparing companies with high and low ESG ratings, hereby being in line with the previous observations as well (H_0 holds).

	Alpha	Rm-Rf	SMB	HML	RMW	CMA	R2
Env. Top	-3.712 (0.27)	1.109*** (0.00)	0.343 (0.33)	0.564** (0.03)	-0.328 (0.36)	-0.067 (0.86)	0.945
Env. Bottom	-3.206 (0.55)	1.130*** (0.00)	0.350 (0.54)	0.688* (0.09)	0.174 (0.76)	0.062 (0.92)	0.841
Soc. Top	-3.359 (0.46)	1.205*** (0.00)	0.854* (0.10)	0.054 (0.86)	-0.015 (0.98)	-0.079 (0.88)	0.906
Soc. Bottom	-4.069 (0.46)	1.330*** (0.00)	-0.062 (0.91)	0.861** (0.05)	0.206 (0.73)	0.428 (0.49)	0.863
Gov. Top	-3.001 (0.46)	1.136*** (0.00)	0.512 (0.24)	0.151 (0.60)	-0.253 (0.56)	0.066 (0.88)	0.920
Gov. Bottom	-10.726** (0.03)	1.166*** (0.00)	1.008** (0.05)	0.267 (0.41)	0.457 (0.35)	0.053 (0.92)	0.890
ESG Top	-2.358 (0.61)	1.099*** (0.00)	0.602 (0.24)	0.477 (0.17)	-0.094 (0.85)	-0.115 (0.83)	0.888
ESG Bottom	-2.175 (0.76)	1.043** (0.02)	0.325 (0.67)	0.930* (0.09)	-0.318 (0.68)	0.018 (0.98)	0.782

Table 6.) The OLS regression results over the whole sample period between 2002 and 2017, implemented with the Fama and French (2015) five-factor model. Alpha (α) signifies the estimated coefficient, that is, the abnormal returns that cannot be explained by the five factors. Rm-Rf, SMB, HML, RMW, and CMA signifies the factor loadings of the beta coefficients. R2 indicates the goodness-of-fit. The p-values are placed below the results, inside the parenthesis. *, **, and *** indicate the statistical significance at the 10%, 5%, and 1% level, respectively.

When investigating the above presented OLS regression results, implemented with the Fama and French (2015) five-factor model, under the H_1 : “incorporating high ESG criteria leads to *positive* abnormal stock returns in the financial sector”, there are again several observations that seem to reject this hypothesis. To begin with, as in the results implemented with CAPM and the three-factor model, also in these results all the portfolios grouped on the highest ESG criteria indicate that the respective dimensions are negatively related to the stock returns. In addition, none of these best-in-class portfolios are statistically significant at any level. Furthermore, the portfolios sorted on the lowest ESG criteria, i.e. the worst-in-class portfolios, also indicate that the respective dimensions are negatively related to the stock returns. Of these, only the “Governance Bottom” portfolio is statistically significant at the 5% level. Therefore, based on these observations, the H_1 can be *rejected*, as it is yet again obvious that incorporating high

ESG criteria does *not* lead to *positive* abnormal stock returns in the financial sector when implemented with these methods and data.

Proceeding with the analysis of the five beta coefficient factors used in the regression model, i.e. $R_m - R_{f_t}$, SMB, HML, RMW, and CMA, certain findings can be made. These findings are rather consistent with the previous results where the three-factor model was used. Firstly, the market factor, i.e. $R_m - R_{f_t}$, is in all cases positively related to the excess stock returns and highly statistically significant at the 1% significance level, except in the case of “ESG Bottom” portfolio where it is statistically significant at the 5% level. This indicates that the constructed best-in-class and worst-in-class portfolios’ excess returns, $R_{it} - R_{Ft}$, are for the most part driven by the markets. Furthermore, when examining the size (SMB_t) and value (HML_t) factors, minor consistencies can be detected. For instance, it can be noticed that the excess returns of the “Social Top” and “Governance Bottom” portfolios are positively related with the size factor, at the statistical significance level of 10% and 5%, respectively. This finding illustrates that small financial companies tend to have higher excess stock returns compared to large ones, if only among these two particular portfolios. In addition, as previously stated, it also seems that the excess returns of some worst-in-class portfolios are driven by the value factor. It can be detected that the excess returns of all the worst-in-class portfolios (excluding Governance) are positively related with the value factor, at the 10% or 5% levels of statistical significance. This demonstrates that the worst-in-class portfolios’ excess returns are driven by the value factor to some extent, expressing that value companies are outperforming growth companies in the financial sector, if only among these specific ESG worst-in-class portfolios. Moreover, the RMW_t (profitability) and the CMA_t (investment) factors express no statistical significance whatsoever. Therefore, it can be stated that using the Fama and French five-factor model is rather superfluous in the context of this thesis, as the Fama and French three-factor model is able to execute as relevant results altogether. For example, the difference between R-squared of “Environmental Top” portfolios between the three-factor and five-factor model is marginal (0.94 vs. 0.945), hereby proving the previously mentioned statement of relevance.

Moreover, when investigating the OLS regression results under the H_2 : “incorporating high ESG criteria leads to *negative* abnormal stock returns in the financial sector”, there are observations that appear to *reject* also this hypothesis. To begin with, all the created best-in-class portfolios indicate that the respective dimensions are, as previously mentioned, negatively related to the stock returns. However, none of the observed abnormal returns are statistically significant at any level. Therefore, the findings indicate that the null hypothesis (H_0) gets *accepted*, as it is evident that the abnormal returns of the created portfolios are not statistically significant.

Furthermore, as already mentioned in the first chapter of the thesis, the null hypothesis will *hold* if statistical significance does not exist in the data sample. In other words, there might occur some abnormal stock returns, yet the probability value, i.e. p-value, of the statistical model indicates that the results are not statistically significant. The findings implemented with the Fama and French (2015) five-factor model are rather controversial when comparing to the findings implemented with the Fama and French (1993) three-factor model. Altogether, the three-factor model indicates that H_2 holds, whereas the five-factor model suggests that the null hypothesis (H_0) gets accepted. This means that according to the five-factor model incorporating high ESG criteria leads to neither positive nor negative abnormal stock returns in the financial sector.

Concluding, these findings are in line with one of the more recent studies of Belghitar, Clark, and Deshmukh (2014), whom suggest that there is no significant difference between the performance of socially responsible investments and conventional investments. By using previous research and empirical mean-variance evidence, the authors find the results to be truly insignificant. Other previous studies that find the same phenomenon of insignificance are from Hamilton, Jo, and Statman (1993) as well as from Bauer, Koedijk, and Otten (2005). Moreover, as mentioned in the interpretation of previous results, also Halbritter et al. (2015) demonstrate that high ESG portfolios do not yield any positive or negative abnormal returns, thus accepting the null hypothesis (H_0) as well.

4.2. Post-crisis sample period

		Alpha	Rm-Rf	SMB	HML	R2
Env. Top	(1)	-9.815 (0.11)	1.263*** (0.01)	-0.050 (0.91)	0.512 (0.13)	0.899
	(2)	-11.280* (0.07)	1.340*** (0.00)			0.782
Env. Bottom	(1)	-2.312 (0.70)	1.459*** (0.01)	0.202 (0.72)	0.912** (0.04)	0.924
	(2)	-6.312 (0.49)	1.704*** (0.01)			0.682
Soc. Top	(1)	-2.773 (0.62)	0.951** (0.04)	0.179 (0.73)	0.082 (0.79)	0.796
	(2)	-3.900 (0.38)	1.032*** (0.00)			0.774
Soc. Bottom	(1)	-13.451 (0.14)	2.085*** (0.01)	-0.310 (0.68)	0.914* (0.10)	0.902
	(2)	-15.015 (0.13)	2.142*** (0.00)			0.774
Gov. Top	(1)	-3.319 (0.58)	1.059** (0.03)	0.025 (0.96)	0.135 (0.69)	0.795
	(2)	-3.886 (0.40)	1.093*** (0.00)			0.780
Gov. Bottom	(1)	-15.173*** (0.01)	1.509*** (0.00)	0.145 (0.62)	0.656** (0.02)	0.975
	(2)	-18.052** (0.02)	1.686*** (0.00)			0.821
ESG Top	(1)	-6.677 (0.33)	1.196** (0.03)	-0.064 (0.92)	0.323 (0.40)	0.807
	(2)	-7.447 (0.20)	1.232*** (0.00)			0.757
ESG Bottom	(1)	-18.279 (0.16)	2.008** (0.04)	0.744 (0.50)	0.390 (0.56)	0.835
	(2)	-23.133* (0.06)	2.357*** (0.00)			0.757

Table 7.) The OLS regression results over the post-crisis sample period between 2010 and 2017, implemented with (1) the Fama and French (1993) three-factor model as well as with (2) the CAPM. Alpha (α) signifies the estimated coefficient, that is, the abnormal returns that cannot be explained by the beta coefficient factors. Thus, Rm-Rf, SMB, and HML signifies the factor loadings of the beta coefficients. R2 indicates the goodness-of-fit. The p-values are placed below the results, inside the parenthesis. *, **, and *** indicate the statistical significance at the 10%, 5%, and 1% level, respectively.

When examining the OLS regression results from the post-crisis sample period, implemented with the CAPM and Fama and French (1993) three-factor model, under the H_1 : “incorporating high ESG criteria leads to *positive* abnormal stock returns in the financial sector”, there seems to be also various findings that disproves this hypothesis. Firstly, all the best-in-class (i.e. “Top”) portfolios indicate that the respective dimensions are negatively related to stock performance, yet only “Environmental Top” portfolio is statistically significant at the 10% level implemented with the CAMP. Less surprisingly, the worst-in-class (i.e. “Bottom”) portfolios also show that the respective dimensions are negatively related to stock performance, of which the “Governance Bottom” portfolio is statistically significant at the 1% level implemented with the Fama and French three-factor model. In conclusion, based on these findings H_1 can be *rejected*, as it is evident that incorporating high ESG criteria does *not* lead to *positive* abnormal stock returns in the financial sector when implemented with these methods and data.

Continuing, when investigating the post-crisis sample’s OLS regression results under the H_2 : “incorporating high ESG criteria leads to *negative* abnormal stock returns in the financial sector”, one cannot obtain clear findings that would accept this hypothesis. All the best-in-class portfolios indicate that the respective dimensions are negatively related to the stock returns, but the results are not statistically significant. Therefore, these findings indicate that the null hypothesis, H_0 , is accepted, as incorporating high ESG clearly criteria leads to neither positive nor negative abnormal stock returns in the financial sector. However, when examining the results more closely, it is evident that some of the worst-in-class portfolios seem to generate statistically significant negative alpha. Implemented with the Fama and French three-factor model, the “Governance Bottom” portfolio seems to generate annual negative alpha of -15.17% at the 1% significance level. Nevertheless, the purpose of this thesis is to investigate how high ESG scores affect stock performance in the financial sector, and thus these previously mentioned results are not significant by nature.

	Alpha	Rm-Rf	SMB	HML	RMW	CMA	R2
Env. Top	-7.464 (0.46)	1.159 (0.12)	-0.114 (0.90)	0.469 (0.41)	-0.458 (0.62)	0.044 (0.97)	0.915
Env. Bottom	6.914 (0.25)	1.031** (0.05)	0.620 (0.27)	1.108** (0.04)	-1.251* (0.09)	-0.841 (0.22)	0.988
Soc. Top	-5.355 (0.63)	1.068 (0.17)	0.128 (0.90)	0.063 (0.91)	0.404 (0.70)	0.134 (0.91)	0.815
Soc. Bottom	-3.097 (0.77)	1.602* (0.09)	0.230 (0.82)	1.172 (0.15)	-1.347 (0.27)	-1.051 (0.42)	0.959
Gov. Top	-0.597 (0.95)	0.948 (0.16)	-0.283 (0.75)	-0.043 (0.93)	-0.720 (0.44)	0.406 (0.70)	0.881
Gov. Bottom	-10.300** (0.04)	1.281*** (0.01)	0.414 (0.18)	0.786** (0.02)	-0.622* (0.09)	-0.517 (0.17)	0.996
ESG Top	-1.542 (0.90)	0.957 (0.23)	0.180 (0.87)	0.438 (0.52)	-0.687 (0.56)	-0.486 (0.72)	0.848
ESG Bottom	-4.165 (0.70)	1.376 (0.12)	0.759 (0.49)	0.349 (0.57)	-2.423 (0.11)	-0.339 (0.78)	0.965

Table 8.) The OLS regression results over the post-crisis sample period between 2010 and 2017, implemented with the Fama and French (2015) five-factor model. Alpha (α) signifies the estimated coefficient, that is, the abnormal returns that cannot be explained by the five factors. Rm-Rf, SMB, HML, RMW, and CMA signifies the factor loadings of the beta coefficients. R2 indicates the goodness-of-fit. The p-values are placed below the results, inside the parenthesis. *, **, and *** indicate the statistical significance at the 10%, 5%, and 1% level, respectively.

When observing the above OLS regression results, implemented with the Fama and French (2015) five-factor model, under the H_1 : “incorporating high ESG criteria leads to *positive* abnormal stock returns in the financial sector”, there seems to yet again be several notions that seem to reject this hypothesis. First of all, as earlier with the whole sample, also in this latter sample period, all the best-in-class portfolios indicate that the respective dimensions are negatively related to the stock returns. In addition, none of these best-in-class portfolios are statistically significant at any level, indicating that the H_1 can be *rejected*. It is yet again obvious that incorporating high ESG criteria does *not* lead to *positive* abnormal stock returns in the financial sector when implemented with these methods and data.

Furthermore, when investigating the OLS regression results under the H_2 : “incorporating high ESG criteria leads to *negative* abnormal stock returns in the financial sector”, there are findings that *reject* also this hypothesis. All the created best-in-class portfolios indicate that the respective dimensions are, as previously mentioned, negatively related to the stock returns, however none of these alphas being statistically significant at any level. Therefore, also these findings indicate that the null hypothesis (H_0) gets *accepted*, as it is evident that the created portfolios’ abnormal returns are not statistically significant. As with the previous methods, also with the five-factor model, the “Governance Bottom” portfolio seems to generate negative alpha of -10.30% annually at the 5% level. Yet again, the purpose of this thesis is to investigate how high ESG scores affect stock performance, and therefore these previously results are not significant by nature.

4.3. Summary of the results

This chapter summarizes as well as provides the final analysis of the OLS regression results executed in this thesis. To review, the empirical results were implemented with three different regression models: the CAPM, the Fama and French (1993) three-factor model, and the Fama and French (2015) five-factor model. Two separate sample periods were used: a whole sample period between 2002 and 2017 as well as a post-crisis sample period between 2010 and 2017. Furthermore, the investigated portfolios were created using the best-in-class (worst-in-class) approach, i.e. screening 20% of the best (worst) performing financial companies listed in the NYSE between 2002 and 2017 by their individual and combined ESG criteria.

		Alpha	Rm-Rf	SMB	HML	RMW	CMA	R2
Env. Top	(1)	-3.712 (0.27)	1.109*** (0.00)	0.343 (0.33)	0.564** (0.03)	-0.328 (0.36)	-0.067 (0.86)	0.945
	(2)	-5.940*** (0.01)	1.222*** (0.00)	0.388 (0.14)	0.542** (0.02)			0.940
	(3)	-4.727 (0.13)	1.272*** (0.00)					0.849
Env. Bottom	(1)	-3.206 (0.55)	1.130*** (0.00)	0.350 (0.54)	0.688* (0.09)	0.174 (0.76)	0.062 (0.92)	0.841
	(2)	-2.006 (0.53)	1.067*** (0.00)	0.343 (0.39)	0.707** (0.04)			0.840
	(3)	-0.685 (0.86)	1.107*** (0.00)					0.703
Soc. Top	(1)	-3.359 (0.46)	1.205*** (0.00)	0.854* (0.10)	0.054 (0.86)	-0.015 (0.98)	-0.079 (0.88)	0.906
	(2)	-3.503 (0.21)	1.217*** (0.00)	0.809** (0.03)	0.033 (0.90)			0.906
	(3)	-2.175 (0.50)	1.338*** (0.00)					0.842
Soc. Bottom	(1)	-4.069 (0.46)	1.330*** (0.00)	-0.062 (0.91)	0.861** (0.05)	0.206 (0.73)	0.428 (0.49)	0.863
	(2)	-2.456 (0.47)	1.220*** (0.00)	0.147 (0.72)	0.979*** (0.01)			0.855
	(3)	-1.150 (0.80)	1.227*** (0.00)					0.696
Gov. Top	(1)	-3.001 (0.46)	1.136*** (0.00)	0.512 (0.24)	0.151 (0.60)	-0.253 (0.56)	0.066 (0.88)	0.920
	(2)	-4.658* (0.08)	1.211*** (0.00)	0.618* (0.06)	0.166 (0.49)			0.916
	(3)	-3.489 (0.24)	1.301*** (0.00)					0.861
Gov. Bottom	(1)	-10.726** (0.03)	1.166*** (0.00)	1.008** (0.05)	0.267 (0.41)	0.457 (0.35)	0.053 (0.92)	0.890
	(2)	-7.637** (0.02)	1.012*** (0.00)	0.921** (0.02)	0.287 (0.30)			0.880
	(3)	-5.854 (0.14)	1.145*** (0.00)					0.736
ESG Top	(1)	-2.358 (0.61)	1.099*** (0.00)	0.602 (0.24)	0.477 (0.17)	-0.094 (0.85)	-0.115 (0.83)	0.888
	(2)	-3.051 (0.29)	1.142*** (0.00)	0.556 (0.12)	0.445 (0.12)			0.887
	(3)	-1.677 (0.63)	1.218*** (0.00)					0.790
ESG Bottom	(1)	-2.175 (0.76)	1.043** (0.02)	0.325 (0.67)	0.930* (0.09)	-0.318 (0.68)	0.018 (0.98)	0.782
	(2)	-4.291 (0.34)	1.144*** (0.00)	0.419 (0.44)	0.931** (0.04)			0.778
	(3)	-2.602 (0.62)	1.192*** (0.00)					0.606

Table 9.) Summary of the OLS regression results over the whole sample period between 2002 and 2017, implemented with (1) the Fama and French five-factor model, (2) the three-factor model, and (3) the CAPM. Alpha (α) signifies the estimated coefficient, that is, the abnormal returns that cannot be explained by the five factors. Rm-Rf, SMB, HML, RMW, and CMA signifies the factor loadings of the beta coefficients. R2 indicates the goodness-of-fit. The p-values are placed below the results, inside the parenthesis. *, **, and *** indicate the statistical significance at the 10%, 5%, and 1% level, respectively.

		Alpha	Rm-Rf	SMB	HML	RMW	CMA	R2
Env. Top	(1)	-7.464 (0.46)	1.159 (0.12)	-0.114 (0.90)	0.469 (0.41)	-0.458 (0.62)	0.044 (0.97)	0.915
	(2)	-9.815 (0.11)	1.263*** (0.01)	-0.050 (0.91)	0.512 (0.13)			0.899
	(3)	-11.280* (0.07)	1.340*** (0.00)					0.782
Env. Bottom	(1)	6.914 (0.25)	1.031** (0.05)	0.620 (0.27)	1.108** (0.04)	-1.251* (0.09)	-0.841 (0.22)	0.988
	(2)	-2.312 (0.70)	1.459*** (0.01)	0.202 (0.72)	0.912** (0.04)			0.924
	(3)	-6.312 (0.49)	1.704*** (0.01)					0.682
Soc. Top	(1)	-5.355 (0.63)	1.068 (0.17)	0.128 (0.90)	0.063 (0.91)	0.404 (0.70)	0.134 (0.91)	0.815
	(2)	-2.773 (0.62)	0.951** (0.04)	0.179 (0.73)	0.082 (0.79)			0.796
	(3)	-3.900 (0.38)	1.032*** (0.00)					0.774
Soc. Bottom	(1)	-3.097 (0.77)	1.602* (0.09)	0.230 (0.82)	1.172 (0.15)	-1.347 (0.27)	-1.051 (0.42)	0.959
	(2)	-13.451 (0.14)	2.085*** (0.01)	-0.310 (0.68)	0.914* (0.10)			0.902
	(3)	-15.015 (0.13)	2.142*** (0.00)					0.774
Gov. Top	(1)	-0.597 (0.95)	0.948 (0.16)	-0.283 (0.75)	-0.043 (0.93)	-0.720 (0.44)	0.406 (0.70)	0.881
	(2)	-3.319 (0.58)	1.059** (0.03)	0.025 (0.96)	0.135 (0.69)			0.795
	(3)	-3.886 (0.40)	1.093*** (0.00)					0.780
Gov. Bottom	(1)	-10.300** (0.04)	1.281*** (0.01)	0.414 (0.18)	0.786** (0.02)	-0.622* (0.09)	-0.517 (0.17)	0.996
	(2)	-15.173*** (0.01)	1.509*** (0.00)	0.145 (0.62)	0.656** (0.02)			0.975
	(3)	-18.052** (0.02)	1.686*** (0.00)					0.821
ESG Top	(1)	-1.542 (0.90)	0.957 (0.23)	0.180 (0.87)	0.438 (0.52)	-0.687 (0.56)	-0.486 (0.72)	0.848
	(2)	-6.677 (0.33)	1.196** (0.03)	-0.064 (0.92)	0.323 (0.40)			0.807
	(3)	-7.447 (0.20)	1.232*** (0.00)					0.757
ESG Bottom	(1)	-4.165 (0.70)	1.376 (0.12)	0.759 (0.49)	0.349 (0.57)	-2.423 (0.11)	-0.339 (0.78)	0.965
	(2)	-18.279 (0.16)	2.008** (0.04)	0.744 (0.50)	0.390 (0.56)			0.835
	(3)	-23.133* (0.06)	2.357*** (0.00)					0.757

Table 10.) Summary of the OLS regression results over the post-crisis sample period between 2010 and 2017, implemented with (1) the Fama and French five-factor model, (2) the three-factor model, and (3) the CAPM. Alpha (α) signifies the estimated coefficient, that is, the abnormal returns that cannot be explained by the five factors. Rm-Rf, SMB, HML, RMW, and CMA signifies the factor loadings of the beta coefficients. R2 indicates the goodness-of-fit. The p-values are placed below the results, inside the parenthesis. *, **, and *** indicate the statistical significance at the 10%, 5%, and 1% level, respectively.

The CAPM

As one would expect, the OLS regressions results obtained with the Capital Asset Pricing Model are rather straightforward. There is only one independent variable, i.e. the market factor, in the model that is explaining the dependent variable, i.e. the excess stock returns. Therefore, it seems rather obvious that the market factor is in all portfolios positively related to the stock returns as well as highly statistically significant at the 1% significance level. This simply indicates, that according to the CAPM, the created portfolios' excess stock returns, $R_{it}-R_{Ft}$, are principally driven by the markets.

Furthermore, all the obtained alphas that cannot be in this case explained by the market factor, seem to be negatively related to the excess stock returns. However, none of the alphas are statistically significant, therefore indicating that the null hypothesis, H_0 , gets *accepted*. In other words, there seems to occur some negative abnormal stock returns possibly explained by the ESG criteria, however the probability value of the statistical model indicating that the results are not statistically significant. Hereby, according to the CAPM, incorporating high ESG criteria leads to neither positive nor negative abnormal stock returns in the financial sector.

The Fama and French three-factor model

When observing the OLS regression results implemented with the Fama and French (1993) three-factor model, some additional interesting information can be obtained. To begin with the analysis of the beta coefficient factors used in the three-factor model, i.e. the market factor (R_m-R_f), the size factor (SMB_t), and the value factor (HML_t), additional observations can be made. Firstly, as in the CAPM, also in the three-factor model the market factor is in all cases positively related to the excess stock returns and highly statistically significant at the 1% significance level, thus demonstrating that the portfolios' stock performance is yet again mostly driven by the markets. However, in addition, it appears that the excess returns of some portfolios are also driven by the size and value factors. For instance, over the whole sample period the excess returns of the "Social Top" and the "Governance Top" portfolios are positively related with the size

factor, at the 5% level and 10% level, respectively. This indicates, that among these particular created portfolios, small financial companies tend to have higher excess stock returns than the large ones. Furthermore, it appears that the stock performance of some worst-in-class portfolios are driven by the value factor. All the worst-in-class portfolios, excluding Governance, are positively related with the value factor at the 1% level or 5% level of statistical significance, indicating that in some cases value companies are outperforming growth companies in the financial sector.

Consequently, when examining the results of the whole sample period, the first alternative hypothesis H_1 : “incorporating high ESG criteria leads to positive abnormal stock returns in the financial sector”, can be rejected. All the best-in-class portfolios indicate that the alphas of the respective dimensions are negatively related to the stock performance, and of these portfolios the “Environmental Top” and the “Governance Top” are statistically significant at the 1% level and 10% level, respectively. Hereby, the first alternative hypothesis (H_1) can be *rejected*, as it is obvious that incorporating high ESG criteria does *not* lead to *positive* abnormal stock returns in the financial sector when implemented with these particular methods and data. Furthermore, when investigating the same OLS regression results under the second alternative hypothesis H_2 : “incorporating high ESG criteria leads to negative abnormal stock returns in the financial sector”, there are findings that appear to accept the hypothesis. As mentioned, all the portfolios grouped on the highest ESG criteria indicate that the alphas of the respective dimensions are, in fact, negatively related to the stock performance, and some of these portfolios are even statistically significant. However, when observing the post-crisis sample period, it seems that incorporating high ESG criteria leads to neither positive nor negative abnormal stock returns. Thus, the overall findings implemented with the three-factor model indicate that either the null hypothesis (H_0) or second alternative hypothesis (H_2) gets accepted.

The Fama and French five-factor model

Finally, when investigating the five-factor model’s OLS regression results under the H_1 : “incorporating high ESG criteria leads to *positive* abnormal stock returns in the financial sector”, there are also findings that reject the hypothesis. As in the previous models, also

in the results of the five-factor model, all the best-in-class portfolios indicate that the respective dimensions are negatively related to the stock performance. However, none of these best-in-class portfolios are statistically significant at any level. Furthermore, the worst-in-class portfolios also indicate that the respective dimensions are negatively related to the stock returns, and of these only the “Governance Bottom” portfolio is statistically significant at the 5% level. Thus, based on these findings, the H_1 can be *rejected*, as it is yet again obvious that incorporating high ESG criteria does *not* lead to *positive* abnormal stock returns in the financial sector when implemented with these particular methods and data.

Consequently, when investigating the OLS regression results under the H_2 : “incorporating high ESG criteria leads to *negative* abnormal stock returns in the financial sector”, there are findings that appear to *reject* also this hypothesis. First of all, as previously mentioned, all the created best-in-class portfolios indicate that the respective dimensions are negatively related to the stock returns – and none of the observed abnormal returns are statistically significant at any level. Therefore, indicating that the null hypothesis (H_0) gets *accepted*. This means that according to the five-factor model, incorporating high ESG criteria leads to neither positive nor negative abnormal stock returns in the financial sector.

Summary of the hypotheses and evidence

The findings implemented with the Fama and French (2015) five-factor model are rather controversial when comparing to the findings implemented with the Fama and French (1993) three-factor model. Altogether, the three-factor model indicates that H_2 holds, whereas the five-factor model suggests that the null hypothesis (H_0) gets accepted. In addition, the CAPM also indicates that H_0 get accepted, as the alphas show no statistical significance whatsoever. Hereby, the following table summarizes all the hypotheses as well as the evidence obtained from the OLS regression analyses.

Hypothesis	Evidence	Support
H₀ : <i>Incorporating high ESG criteria leads to neither positive nor negative abnormal stock returns in the financial sector.</i>	Statistically insignificant alphas obtained from the CAPM as well as from the Fama and French (2015) five-factor model, indicating the null hypothesis holds. However, the Fama and French (1993) three-factor model partly rejects this hypothesis when observing the whole sample period.	Yes / No
H₁ : <i>Incorporating high ESG criteria leads to positive abnormal stock returns in the financial sector.</i>	Statistically significant alphas obtained from the Fama and French (1993) three-factor model that reject this hypothesis over the whole sample period.	No
H₂ : <i>Incorporating high ESG criteria leads to negative abnormal stock returns in the financial sector.</i>	Statistically significant alphas obtained from the Fama and French (1993) three-factor model that accept this hypothesis over the whole sample period. “Environmental Top” and “Governance Top” portfolios seem to create negative alpha.	Yes

Table 11.) Summary of the hypotheses and evidence.

5. CONCLUSION

The purpose of this thesis is to examine whether higher ESG scores lead to abnormal stock returns in the financial sector. The thesis focuses only on the financial markets of United States, more specifically on the banks and financial services companies listed in the New York Stock Exchange over the period between January 2002 and January 2017. Financial sector was selected due to its troublesome reputation in corporate responsibility and ethics altogether. Consequently, this thesis aims to find evidence to the following research question: “How does the incorporation of ESG criteria affect stock performance in the financial sector?”.

What comes to the research question, Auer (2016) quite recently showed that it is still possible to achieve abnormal stock returns with SRI in Europe with portfolios screened on the highest ESG scores. Also earlier academic research, such as Kempf’s and Osthoff’s (2007), state that the incorporation of SRI can lead to high positive abnormal stock returns as simply as following a long-short strategy, positive screening approach, or best-in-class approach. Furthermore, Renneboog et al. (2008) find that investors who are using ESG as an investment criteria tolerate a higher cost and are thus willing to accept suboptimal financial performance. Also, Halbritter et al. (2015) demonstrate that ESG portfolios do not yield any abnormal returns when comparing companies with high and low ESG ratings.

Continuing, the thesis uses ESG and financial data obtained from the Thomson Reuters ASSET4 database and factor data obtained from Kenneth R. French’s (2019) database. The data retrieved from the Thomson Reuters’ database is divided into two parts. The first part consists of the yearly, end-of-the-year, individual and combined ESG scores of all the selected financial companies. The second part of the data comprises of the monthly, end-of-the-month, closing prices of all the same companies, resulting in approximately 35,000 monthly observations, covering exactly 181 months over a 16 year period, thus spanning 193 financial companies that are, or were, part of the NYSE investment pool at the given sample period. Furthermore, all the necessary data needed for the regression

analyses, such as the beta coefficient factors of SMB_t , HML_t , RMW_t , and CMA_t , as well as the risk-free rates of return, are all gathered from the Kenneth R. French's database.

The methodology and empirical framework of this thesis complies largely with Derwall et al. (2005), Kempf et al. (2007), Renneboog et al. (2008), as well as with Halbritter et al. (2015). They all make use of the CAPM, Fama and French (1993) three-factor model, or the Carhart (1997) four-factor model, which are all so-called prior versions of the Fama and French (2015) five-factor model. Moreover, Kempf et al. (2007) as well as Halbritter et al. (2015) use the best-in-class approach in their studies, by screening for example 10%, 20%, and 25% of the best (worst) performing stocks sorted by their ESG criteria. This thesis makes use of the best-in-class (worst-in-class) approach as well by screening 20% of the best (worst) performing financial companies listed in the NYSE between 2002 and 2017 by their individual and combined ESG scores. Furthermore, the OLS regression results are divided into two separate sample periods (whole sample and post-crisis sample) and implemented by using three different regression models: the CAPM, the Fama and French (1993) three-factor model, and the Fama and French (2015) five-factor model. What comes to the data and methods, both Thomson Reuters ASSET4 as well as Fama and French (2015) five-factor model are very rarely used in the academic literature of ESG, CSR, and SRI. Hereby, this thesis will also evaluate the usability of them both in the environment of corporate responsibility.

This thesis finds that incorporating higher ESG criteria leads to either negative abnormal stock returns, or does not have an effect at all. According to the regression results obtained with the Fama and French (1993) three-factor model, 20% of the best performing financial companies ranked by their "Environmental" scores, generate abnormal stock returns of -5.94% annually over the whole sample period, whereas the top financial companies ranked by their "Governance" scores generate annual alpha of -4.66%. This indicates that the high ESG scored financial companies underperform the benchmark, i.e. the risk-free investment. These results are in line with the paper of Renneboog et al. (2008), whom find that investors who are using ESG as an investment criteria tolerate a higher cost and are thus accepting inferior stock returns and financial performance overall. Also, Halbritter and Dorfleitner (2015) demonstrate that ESG portfolios do not

yield any abnormal returns when comparing companies with high and low ESG ratings, hereby being in line with the previous observations as well.

Furthermore, the CAPM and the Fama and French (2015) five-factor models provide statistically insignificant alphas, thus indicating that high ESG scored financial companies neither under- or overperform the risk-free investment. These findings are in line with one of the more recent studies of Belghitar et al. (2014), whom suggest that there is no significant difference between the performance of socially responsible investments and conventional investments. Other previous studies that find the results to be truly insignificant are from Hamilton et al. (1993) as well as from Bauer et al. (2005).

Consequently, in order to provide a final answer to the question whether high ESG scored financial companies overperform, underperform, or does neither, the results clearly suggest that they either *underperform* or *does neither*. Depending on the asset pricing model and data sample used in the analyses, the evidence provide either statistically insignificant alphas, or alternatively state that high ESG scored companies are experiencing negative annual alphas.

To discuss, this underperformance might be present due to the fact that responsible financial companies are tolerating higher costs and are thus willing to accept suboptimal financial performance. It could be that in the financial sector profits are rather ruthlessly considered plainly as profits regardless of the corporate actions. Meaning that if the incorporation of Environmental criteria does not yield any straightforward and countable financial benefits, they are then more likely considered as expenses. In addition, to give an example, the benefits from good corporate governance and the disadvantages from bad governance would be much easier to estimate compared to the possible financial benefits achieved from environmental actions. Therefore, it might just be that the true utility achieved from the implementation of ESG, and especially from Environmental dimensions, could just be too difficult to measure, as there is not any distinct way to estimate the financial benefits achieved from it.

Even though the thesis' regression results are chiefly supported by earlier academic literature, there are still some limitations that need to be taken into consideration. First of all, the quality of the financial data acquired from the Thomson Reuters database was not the best possible, as a rather significant amount of closing price data was not accessible. Secondly, the ESG data gathered from the same database was not the best quality either, as some of the ESG ratings were accessible only within the last decade or so. Overall, one needs to remember that the quality of the output is principally driven by the quality of the input. Nonetheless, if there would be a better access to data, these limitations could be corrected altogether.

What comes to the future research on the topic of ESG, it could be insightful to focus on different markets, e.g. Europe or Asia, or to totally another industries. Europe and Asia as data samples could provide significantly different results and focusing on specific industries could provide additional valuable insights as well. Furthermore, employing specific sample periods could be appropriate by concentrating for example on the crisis periods. The main focus could be on the U.S. financial crisis or alternatively on the Eurozone crisis. Furthermore, employing different data, methods, and asset pricing models could be meaningful as well, and implementing different portfolio construction processes could provide interesting results. For example, creating "sin stock" portfolios could offer some interesting insights on the topic altogether.

LIST OF REFERENCES

- Adler, T., & Kritzman, M. (2008). The cost of socially responsible investing. *Journal of Portfolio Management*, 35(1), 52.
- Auer, B. R. (2016). Do socially responsible investment policies add or destroy European stock portfolio value?. *Journal of business ethics*, 135(2), 381-397.
- Bauer, R., Koedijk, K., & Otten, R. (2005). International evidence on ethical mutual fund performance and investment style. *Journal of Banking & Finance*, 29(7), 1751-1767.
- Bebchuk, L. A., Cohen, A., & Wang, C. C. (2013). Learning and the disappearing association between governance and returns. *Journal of financial economics*, 108(2), 323-348.
- Belghitar, Y., Clark, E., & Deshmukh, N. (2014). Does it pay to be ethical? Evidence from the FTSE4Good. *Journal of Banking & Finance*, 47, 54-62.
- Bodie, Z. (2014). *Investments* (10th global ed.). New York: McGraw Hill Education.
- Borgers, A., Derwall, J., Koedijk, K., & Ter Horst, J. (2013). Stakeholder relations and stock returns: On errors in investors' expectations and learning. *Journal of Empirical Finance*, 22, 159-175.
- Brzeszczyński, J., & McIntosh, G. (2014). Performance of portfolios composed of British SRI stocks. *Journal of business ethics*, 120(3), 335-362.
- Crane, A., Matten, D., & Spence, L. (Eds.). (2019). *Corporate social responsibility: Readings and cases in a global context*. Routledge.

- Crowther, D., & Seifi, S. (Eds.). (2018). *Redefining Corporate Social Responsibility*. Emerald Group Publishing.
- Dahlsrud, A. (2008). How corporate social responsibility is defined: an analysis of 37 definitions. *Corporate social responsibility and environmental management*, 15(1), 1-13.
- Davis, K. (1960). Can business afford to ignore corporate social responsibilities. *California management review*, 76.
- Derwall, J., Guenster, N., Bauer, R., & Koedijk, K. (2005). The eco-efficiency premium puzzle. *Financial Analysts Journal*, 61(2), 51-63.
- Eccles, N. S., & Viviers, S. (2011). The origins and meanings of names describing investment practices that integrate a consideration of ESG issues in the academic literature. *Journal of business ethics*, 104(3), 389-402.
- Epstein, M. J. (2018). *Making sustainability work: Best practices in managing and measuring corporate social, environmental and economic impacts*. Routledge.
- Fabozzi, F. J., Ma, K. C., & Oliphant, B. J. (2008). Sin stock returns. *Journal of Portfolio Management*, 35(1), 82.
- Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, 25(2), 383-417.
- Fama, E. F. (1991). Efficient capital markets: II. *The journal of finance*, 46(5), 1575-1617.
- Fama, E. F., & French, K. R. (1992). The Cross-section of Expected Stock Returns. *The Journal of Finance*, 47(2), 427-465.

- Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, 33(1), 3-56.
- Fama, E. F., & French, K. R. (2004). The capital asset pricing model: Theory and evidence. *Journal of economic perspectives*, 18(3), 25-46.
- Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of financial economics*, 116(1), 1-22.
- Fama, E. F., & French, K. R. (2018). Choosing factors. *Journal of Financial Economics*, 128(2), 234-252.
- Garriga, E., & Melé, D. (2004). Corporate social responsibility theories: Mapping the territory. *Journal of business ethics*, 53(1-2), 51-71.
- Gladwin, T. N., Kennelly, J. J., & Krause, T. S. (1995). Shifting paradigms for sustainable development: Implications for management theory and research. *Academy of management Review*, 20(4), 874-907.
- Halbritter, G., & Dorfleitner, G. (2015). The wages of social responsibility—where are they? A critical review of ESG investing. *Review of Financial Economics*, 26, 25-35.
- Hamilton, S., Jo, H., & Statman, M. (1993). Doing well while doing good? The investment performance of socially responsible mutual funds. *Financial Analysts Journal*, 49(6), 62-66.
- Hickman, K. A., Teets, W.R., & John J. Kohls (1999). Social investing and modern portfolio theory. *American Business Review* 17:1, 72-78.
- Ho, V. (2016). Risk-Related Activism: The Business Case for Monitoring Nonfinancial Risk. *Journal of Corporation Law*, 41(3), pp. 647-705.

- Hong, H., & Kacperczyk, M. (2009). The price of sin: The effects of social norms on markets. *Journal of Financial Economics*, 93(1), 15-36.
- Humphrey, J. E., & Tan, D. T. (2014). Does it really hurt to be responsible?. *Journal of business ethics*, 122(3), 375-386
- Kempf, A., & Osthoff, P. (2007). The effect of socially responsible investing on portfolio performance. *European Financial Management*, 13(5), 908-922.
- Kenneth R. French (2019). Data Library. Developed Market Factors and Returns [online]. [cited 15.04.2019]. Available from World Wide Web: <URL: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html>
- Lins, K. V., Servaes, H., & Tamayo, A. (2017). Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis. *The Journal of Finance*, 72(4), 1785-1824.
- Lintner, J. (1965). The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *The Review of Economics and Statistics*, 47(1), pp. 13-37.
- Markowitz, H. (1952). Portfolio selection. *The journal of finance*, 7(1), 77-91.
- Matten, D., Crane, A., & Chapple, W. (2003). Behind the mask: Revealing the true face of corporate citizenship. *Journal of business ethics*, 45(1-2), 109-120.
- Mossin, J. (1966). Equilibrium in a Capital Asset Market. *Econometrica*, 34(4), p. 768.
- New York Stock Exchange (2019). Market Data: NYSE Composite Index [online]. [cited 02.10.2019]. Available from World Wide Web: <URL: <https://www.nyse.com/quote/index/NYA>>

- Nofsinger, J., & Varma, A. (2014). Socially responsible funds and market crises. *Journal of Banking & Finance*, 48, 180-193.
- Renneboog, L., Ter Horst, J., & Zhang, C. (2008). Socially responsible investments: Institutional aspects, performance, and investor behavior. *Journal of Banking & Finance*, 32(9), 1723-1742.
- Revelli, C., & Viviani, J. L. (2015). Financial performance of socially responsible investing (SRI): what have we learned? A meta-analysis. *Business Ethics: A European Review*, 24(2), 158-185.
- Roll, R. (1977). A Critique of the Asset Pricing Theory's Tests Part I: On Past and Potential Testability of the Theory. *Journal of Financial Economics*, 4(2), 129-176.
- Schueth, S. (2003). Socially responsible investing in the United States. *Journal of business ethics*, 43(3), 189-194.
- Sharpe, W. F. (1964). Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk*. *Journal of Finance*, 19(3), pp. 425-442.
- Thomson Reuters (2019). Thomson Reuters ESG Scores [online]. [cited 15.04.2019]. Available from World Wide Web: <URL: https://www.refinitiv.com/content/dam/marketing/en_us/documents/methodology/esg-scores-methodology.pdf>
- UNPRI (2019a). Brochure: Principles for Responsible Investment [online]. [cited 02.07.2019]. Available from World Wide Web: <URL: <https://www.unpri.org/download?ac=6303>>
- UNPRI (2019b). About the PRI [online]. [cited 02.07.2019]. Available from World Wide Web: <URL: <https://www.unpri.org/pri/about-the-pri>>

US SIF (2018). Report on US Sustainable, Responsible and Impact Investing Trends 2018 [online]. [cited 16.04.2019]. Available from World Wide Web: <URL: <https://www.ussif.org/files/US%20SIF%20Trends%20Report%202018%20Release.pdf>>