Copenhagen Business School
Spring 2020

Empirical Analysis of ESG and Financial Performance

Written by:
Mathias Dyrhol Paulsen, Student No.: S124599
Øystein André Pettersen, Student No.: S125465

Hand-in date:
15.05.2020

Master Thesis

MSc Finance & Strategic Management

Copenhagen Business School 2020

Supervisor:
Thomas Einfeldt

No. of Characters: 200.040
No. of Normal Pages: 96

Abstract

Globally, investors and financial markets are directing increasingly more attention towards responsible investments. The term "responsible investing" is often interpreted as environmental, social and governance (ESG) concerns for investment decisions with a long-term perspective. The concept has become increasingly more relevant among consumers, government, investors and stakeholders. The increased focus is not based on empirically superior relationships between risk and return, but rather on a shift in demand for responsible investments with a more long-term perspective. Previous studies focusing on the relationship between ESG and financial performance are split into three distinctions; positive, neutral and a negative relationship. These three counterparts find theoretical arguments and statistical evidence supporting their results, and a clear conclusion regarding the relationship is yet to be made.

This thesis examines the relationship that has puzzled the academia, with a thorough and critical review of existing literature on ESG investing. The empirical analysis examines portfolios with varying degrees of ESG performance, where the performance has been identified by the companies' respective ESG and controversy score. These numbers are provided by Refinitiv, which is the successor of Asset 4, and are collected through Thomson Reuters Datastream. By application of traditional asset pricing models, namely the CAPM, Fama & French three-factor, Carhart four-factor and Fama & French five-factor model, the return on various portfolios has been controlled against known risk factors. Moreover, both ESG and controversy factors have been developed to study the relationships in greater depth. The results find evidence that implies a negative relationship between high ESG scores and excess returns. However, this result is not evident in the robustness tests, where the portfolios are divided into sub-periods and classified into different portfolio sizes. In contrary to previous findings, the analysis finds evidence that the companies with the absolute lowest ESG scores have a negative excess return. Nevertheless, the negative alpha is not substantially different from zero. There is no evidence in the analysis that can provide an answer to the question of whether or not controversies have any effect on the excess return. The results are more supportive of the literature that implies a negative correlation between increased ESG initiatives and financial performance measured by excess returns. However, the question of whether ESG is priced in by the financial market remains open for further investigation.

Preface

This master thesis represents the end of a two-year master's degree program in Finance & Strategic Management at Copenhagen Business School.

The thesis is a quantitative study of responsible investments, where the main focus lay on the effect of ESG criteria and financial performance. The motivation was to do a quantitative study in practice and get interesting insights on the subject.

The authors would like to thank the supervisor, Thomas Einfeldt, for valuable discussions, suggestions and comments during the process of the thesis. The authors would also like to thank Copenhagen Business School for the supply of valuable resources, mainly the use of Thomson Reuters Datastream, which is gratefully acknowledged.

Contents

1 Introduction

This chapter will focus on the current situation and the relevance of responsible investments and ESG. The motivation and research question are presented at the beginning, followed by a more nuanced explanation of the problem area. Next, the scientific method used, and the research design will be described. Finally, criticism and the research scope are presented.

1.1 Motivation

The last decades have witnessed a growing attention towards the non-financial aspects of companies, emphasizing the relevance of responsible investing. Driven by factors such as climate change, environmental issues have become a prominent topic in modern-day society, challenging global politics to confront various environmental concerns (Bauer and Hann, 2010). International conventions, such as the United Nations Global Compact (UNGC) and the United Nations Principles for Responsible Investment (PRI), provide companies with guidelines for their practices and encourage commitment to foundational responsibilities. These responsibilities could include initiatives for e.g. environmental concerns, labor, human rights and anti-bribery (Della Croce, 2011).

The environmental, social and governance (ESG) performance of a company has transitioned into a focus on the forefront of academic research and in practice. The growth trajectory is demonstrated in recent figures. The Global Sustainable Investment Alliance (GSIA) estimated that the total value of all funds managed according to an ESG mandate was $30.7 trillion in the five major markets in 2018. This is further estimated to reach $50 trillion over the next two decades (Stevens, 2019). Europe accounts for the largest concentration of ESG-mandated assets worldwide, totaling $14.1 trillion (GSIA, 2019). Investors seem to have an increasing focus on socially responsible investments (SRI), which illustrates a more long-term perspective for business- specific decisions. Environmental and social concerns, such as greenhouse gas emission and employment rights, have received increasing attention during the last decades, while scandals, such as those from the companies Enron[1] and WorldCom[2], have highlighted the importance of a transparent governance structure.

[1] US energy-trading company that concealed large losses through its own projects. Filed for bankruptcy in 2001. (Bondarenko, 2019)

[2] US telecommunication company that hide billions of dollars in property accounts. Filed for bankruptcy in 2001. (Kennon, 2019)

The use of ESG scores have become an acknowledged tool in order to measure a company's performance on responsible actions and initiatives. An ESG score is a score calculated on the basis of a company's management in regard to environment, social conditions and corporate governance parameters. In practice, hundreds of different subfactors are measured on both qualitative and quantitative criteria for individual companies, which in turn can be used for investment evaluations of different prospects. Financial data providers, including KLD and Thomson Reuters, report on ESG parameters for individual companies on a global scale, allowing investors to use supplementary financial data in their investment decisions.

As the issue becomes more relevant to investors and stakeholders, who are considering different approaches of investing their resources for the future, the relevance of the financial effects of ESG performance as a research area is increasing. However, due to a somehow contradictory existing literature, empirical studies have not been able to produce consistent results on the subject, leaving an uncovered area that is increasingly important for investment decisions worldwide. Therefore, further studies are needed in order to shed light on the mechanisms and relationship between responsible investments and financial performance. In particular, the literature lacks supporting empirical studies that analyze the financial effects of ESG initiatives and performance by including additional relevant factors such as controversies over a longer period of time.

1.2 Research Question

Investments in environmental and social responsibility initiatives are a common topic among companies, investors and researchers, which have led to a central discussion about the actual impact on a company's financial performance. Since the early 1970s, several thousand studies have been published on how responsible investments affect risk and return. The literature review (chapter 2) refers to various meta-studies that include several thousand studies on how responsible investments affect financial performance. In academic research, both theoretical and empirical studies are split on the question of whether a positive, negative or no relationship between ESG and financial performance exists. To this date, empirical analyses have not reached a clear consensus on the actual effect, and the debate on the effect of ESG and financial performance is still to be continued (Friede et al., 2015).

The objective of this thesis is to collect some elements from previous studies as well as adding new elements in order to gain new and valuable insights on the subject. Previous studies have mainly focused on ESG and its impact on returns, but little attention has been paid to controversies connected to the respective companies. Aouadi and Marsat (2018) argue that ESG and controversies related to ESG are not direct contradictions. By including controversies as a supplement to the widely researched ESG scores, this paper may contribute to new information. This thesis will more closely investigate the effects of ESG and controversies on company returns, and whether today's increasing focus on responsible investments can be defended from an economic perspective.

This study will solely focus on European companies and their ESG performance from 2003 until 2019. The ESG performance is measured by the respective companies' ESG and controversy score, which are collected from Thomson Reuters Datastream. Observations prior to 2003 is not evaluated, as the information regarding responsible initiatives is drastically reduced.

The thesis will examine the following research question:

"What financial effect do ESG and Controversies have on returns for European companies?"

The following sections will explain how the process will be conducted and how the research question will be answered.

1.3 Problem Area

Previous theoretical and empirical studies have been based on a positive, neutral or a negative relationship between corporate responsibility and financial performance. The theories have provided different perspectives on the matter and made the need for empirical studies to substantiate whether the theoretical arguments are universally valid or not. In the literature review, two prominent methodological approaches to the issue are presented. The issue has primarily been investigated by studying the performance of socially responsible funds relative to conventional funds, in addition to more recent studies that have used an alternative long-term approach by forming diversified portfolios of single assets.

The performance of Socially Responsible Investment (SRI) funds are a widespread subject area and show mixed results on the relationship between SRI and financial performance. Statman (2000) finds that SRI funds performed worse than the S&P500 in the period 1990-1998, but similar or better than conventional funds. This study, in addition to the studies made by Bauer, Koedijk and Otten (2005) and Bello (2005), does not find any significant relationship between SRI and financial performance. Climent and Soriano (2011) find that environmental funds underperform compared to conventional funds with similar characteristics. However, it must be borne in mind that the validity of the results is influenced by several different factors that have an impact on the return of the funds. The fund studies also suffer from the fact that their SRI parameters are not uniform and have different criteria in their screening of assets. Hence, the studies examine relationships with different subjective perceptions and is not based on unambiguous criteria for responsibility.

The obvious weaknesses of the fund studies presented above are avoided when longer-term studies are conducted. In this case, portfolios of individual assets are compiled according to selected and specified ESG criteria and are not influenced by fund-specific factors that affect the results. However, these types of studies are significantly influenced by the chosen ESG criteria (Dorfleitner et al., 2014), and in particular the choice of the data provider. Berg, Kölbel and Rigobon (2019) analyze five prominent rating agencies and find that measurement divergence explains more than 50 percent of the overall divergence. They also detect a "rater effect", which means that the rating agencies evaluate the ESG categories differently and use pre-defined assessment criteria that differ between the respective data providers. For instance, if a shipping company receives a poor grade on the environmental factor because of pollution, one rating agency could give them an overall low ESG score, while another could give them a high ESG score due to a strong governance structure. This shows that studies based on data from different ESG rating providers could cause the individual results to differ to a greater or lesser extent.

Recent financial theory argues that investors have a utility function that comprises more than just a trade-off between risk and return. Investors also incorporate social values and personal opinions and beliefs into their investment decision (Bollen, 2007). Bollen (2007) finds that socially responsible funds have a lower monthly volatility in cash withdrawals from investors compared to conventional funds. This could indicate that investors are more willing to keep their money in socially responsible funds that includes companies with a long-term and sustainable strategy. This theoretical perspective attempts to explain why companies experience increasing pressure to be socially responsible and limit

their negative externalities. As these developments have evolved, the subject has become increasingly relevant to pension funds and private investors, with respect to setting goals for their portfolio allocation and contribute to future sustainability. This trend contrasts with classical economic theory of investment decisions. The classical mindset was formed by Markowitz (1952), where investors are basing their investment decision on maximizing their utility, or risk-return ratio. One of the concepts concerns the choice of a portfolio where the investor maximizes the discounted value of future returns, i.e. "expected" or "anticipated" returns are discounted. This approach still has a significant influence on investment decisions, but other factors such as the environment and social responsibility are increasingly being taken into account when trying to explain investor behavior (Bollen, 2007).

Several different theorists have argued that sustainable investments have a negative impact on company value. Friedman and Friedman (1962) address that employing a CSR policy is not a core task for the management, and may also result in cash flows being wasted on unprofitable projects. This implies that if corporate social responsibility represents a transfer of wealth from shareholders to stakeholders, socially responsible companies are likely to produce lower equity returns than less socially responsible companies. Conversely, this means that there is a smaller incentive to make investments in sustainable projects or companies. Walley and Whitehead (1994) argue that environmental standards can result in positive consequences for the overall economy, but that these standards will be expensive to implement and will cause problems for individual companies that don't have the capacity to uphold them. Furthermore, they state that these standards will result in higher product prices, lower profitability and competitive disadvantages.

On the other hand, there are a number of theoretical counterparts which argue that socially responsible firms should have better performance. This includes, among other elements, that they attract more qualified employees (Moskowitz, 1972) and that they have a better relationship with stakeholders (Freeman and Reed, 1983). Porter et al. (1995) argue that a firm's investment in ESG initiatives may result in a competitive advantage because of less pollution, improved quality and a more cost-effective use of resources. They argue that socially responsible companies may report significantly higher revenues than less socially responsible companies. Derwall et al. 2005 state that no significant effect on returns can be explained by socially responsible investments, as opposed to theories that argue of either positive or negative effects. The effective market hypothesis and the capital asset pricing model (CAPM), which assumes that the return on an investment is proportional to the

associated risk, contributes to the statement that there should not be a significant difference in the returns of companies that are more or less socially responsible (Derwall et al., 2005).

Previous research on the long-term relationship between ESG and financial performance do not illustrate a consistent answer. Several studies, such as (Russo and Fouts 1997), (Dowell et al. 2000), (Bollen, 2007), (Derwall et al. 2005), (Friede, Busch and Bassen 2015) find a positive correlation. Studies have shown that increased investments in ESG results in lower cost of capital (Sharfman and Fernando, 2008) and lower cost of equity (El Ghoul et al. 2011). Other studies, by contrast, find a neutral relationship (Cohen et al., 1997), (Galema et al., 2008), (Halbritter and Dorfleitner 2015) or negative relationship (Waddock and Graves, 1997), (Berman et al., 1999), (Bauer et al., 2005), (Oberndorfer et al., 2013), (Zhang, 2017).

The majority of previous portfolio studies have used stock returns or financial ratios as proxies for the financial performance. An overweight of the studies that investigates the relationship between ESG and financial performance find a positive correlation using this approach. As various approaches seem to obtain diverging results, there is no clear evidence that it exists a positive relationship between ESG and financial performance. How methods affect results will be further discussed in the literature review.

1.4 Scientific Method

This thesis attributes a positivistic and epistemological approach for the processing of knowledge and information, which means that only observable phenomena lead to the validation of data and the problem formulation. The investigator (the authors) and the investigated object (ESG and financial performance) are assumed to be independent entities, where the investigator can study the object without influencing it to a pre-defined destination nor being influenced by previous research, i.e. processing the data with an objective approach (Guba and Lincoln, 1994). The steps that are used to handle the data and the conclusions that have been made are clarified so the reviewer has the possibility to judge the results, either by accepting or rejecting the contributions to science and future research (NAS, 2009). By thoroughly explaining terms and data, the aim is to provide a high degree of reliability and the opportunity to validate the quality of the data, in order to make it possible to replicate the data for further research.

The results from the analysis and sub-conclusions drawn throughout the thesis will uncover general relationships that can be observed in the real world in accordance with the positivistic approach. Furthermore, the data collection will be conducted in accordance with the research method chosen in order to make an objective data collection. The thesis will use a deductive approach where the gathered data is structured for the purpose of optimizing its contribution to test the research question presented.

The research question is prepared by reviewing the most influential literature on responsible investments and ESG performance. Section 1.6 is dedicated to explanations on how the most influential literature and high-quality studies are evaluated. The literature review is meant to give nuanced and in-depth insights into the context, study methods and theories on the investigated subject area. The problem formulation is constructed based upon the results provided by the literature review. Furthermore, the problem formulation is formulated on the basis of similar approaches to the problem from the literature, where the testing will use the most prominent and tested methods in the literature.

All equities, indexes and risk targets that have been used in this thesis can be observed and collected objectively in the financial market. As an objective perspective is emphasized, the assembled data ensures reliability and validity to the results of the regression analysis. The aim is that the test results do not depend on the authors' views or work on the data (Guba and Lincoln, 1994).

1.5 Research Design

The thesis will follow a stringent structure that is consistent with the scientific method. The research design will provide a theoretical and empirical foundation for further quantitative processing, analysis and discussion of the research question.

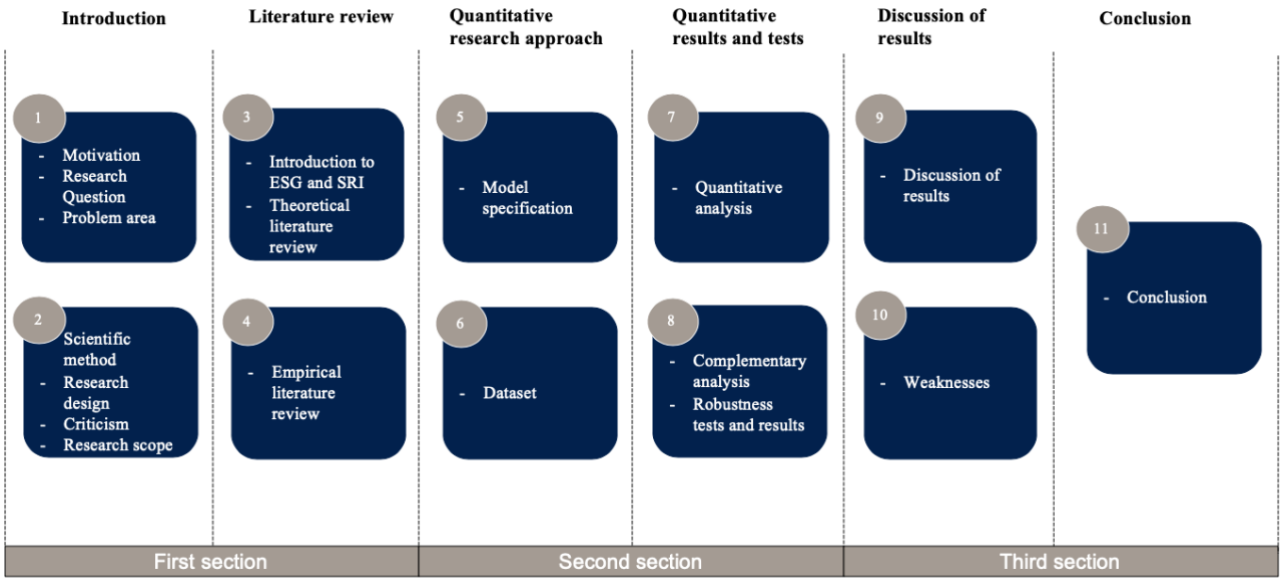


Figure 1: Thesis structure

The first part of the thesis will introduce and describe the subject area that is investigated. Moreover, the relevance of the subject and its impact on the financial market will be elaborated upon. The research method will be explained to ensure the reader receives relevant insights into the scientific approach of how the authors have processed the information and data. Furthermore, ESG and its conceptual background will be described. The reader will be informed about the developments on responsible investments and ESG initiatives in relation to financial performance throughout the years.

With an understanding of the interaction between ESG and financial performance, the thesis will continue by presenting a detailed literature review based on the most influential studies on the subject. This section will describe the theoretical arguments that imply positive, neutral and negative relationships between ESG and financial performance. On the basis of the theoretical arguments, the thesis will provide a literature review of empirical studies that have investigated the subject based on the theoretical approaches. Furthermore, the authors will identify potential gaps where this thesis could contribute to new or supporting information to the already existing literature.

The second part will focus on the different models, data and variables that are applied in the analysis. These have been chosen on the basis of the methods suggested from the literature review, in order to identify possible relationships and correlation between ESG and financial performance. The data collection and the processing of this data will be comprehensively explained to ensure reliability. To conclude the second section, complementary analysis and robustness tests will be presented to test the validity of the initial analysis, followed by a summary and discussion part of the analysis.

In the third section, the quantitative results and implications will be discussed as well as the significance of ESG from an investor perspective. It will attempt to answer the research question with an in-depth discussion of the test results. These results will then be included and discussed in comparison to similar empirical studies on the subject. Finally, the thesis will suggest different approaches that can be made on the investigated problem area that are relevant for further research.

1.6 Criticism

In the process of gathering information for this thesis, the emphasis has been on collecting influential articles, scientific journals and other relevant publications to ensure quality and validity of the used material. The focus is to thoroughly explore and study previous research that could be used for further analysis and interpretation. The literature has mainly been collected through CBS Libsearch, which again uses electronic databases such as Springer Link, SSRN, JSTOR, Science Direct, Emerald Insights, Wiley-Blackwell, Wiley-Online and Elsevier. Web search engines such as Google Scholar were also used, in addition to material found through Harvard Business Review. Appropriate key words such as “Responsible investing”, “SRI”, “ESG and financial performance”, “Responsible investing and ESG”, “Ethical investing”, “Corporate environmental performance”, “Corporate social performance” and “CSR performance”, have been entered to locate relevant material. The thesis has used proxies to ensure the quality e.g. number of citations in studies, journals and papers and their respective ranking through SJR[3]. The authors are aware that this is not a complete measure to ensure

[3] Scientific Journal Rankings – SJR. https://www.scimagojr.com/journalrank.php

the quality of the relevant study's influence and quality. A more comprehensive assessment of the studies has thus been produced to ensure the usability of the material used for this thesis.

In compliance with Revelli and Viviani (2015), the publications were controlled in order to avoid overrepresentation, where duplicated studies were excluded. The studies that were found to be the least successful with regard to methodology and results, such as working papers that had a lack of quantitative data, were removed. In addition, studies that did not include statistical data (Hunter and Schmidt, 2004) or did not use financial performance as the dependent variable were excluded. Studies without ESG, SRI or CSR as experimental group were also eliminated. Next, according to Hoepner and McMillan (2009) and Armstrong and Wilkinson (2005;2007), the literature review has been assessed according to 11 assessment criteria. The criteria are illustrated in section 12.1, where the reader will get a deeper insight into the subject area of the study, methodology, reliability, conclusions and implications for further research. By following these 11 assessment criteria, the thesis ensures that the most influential studies have been processed and evaluated.

Revelli and Viviani (2015) define publication bias as the likelihood that only studies that have statistically significant or interesting results, are being published. This means that studies that don't have significant results, or findings that don't contribute to existing research, are less likely to be made available (Song et al., 2000). This could in turn lead to selective publications, where the literature will focus only on significant and positive results for the subject. In this thesis, the aim has been to locate representative studies that show a positive, neutral and negative relationship. This also includes studies that do not show significant results. Both the theoretical and the empirical literature review include a clear context of the results and what the studies were able to find.

After an objective evaluation of the literature according to Hoepner and McMillan (2009) and the study of selection and publication bias of Revelli and Viviani (2015), this thesis will include high-quality studies that focus on ESG and financial performance. It is further assumed that the necessary validity measures of the studies have been taken.

1.7 Research Scope

This thesis will examine the relationship between companies' ESG- and Controversy Scores and financial performance in the European market. ESG will be studied on a broader measure, i.e. the individual pillars (E,S,G) will not be examined individually. The relationship is therefore not investigated on a global level, but limited to Europe. Hence, the results presented are not necessarily valid in other geographical locations. Additionally, the study is delimited to include only listed companies with published data from Thomson Reuters Datastream.

Throughout the paper, new and less-known relevant theories will be elaborated upon. However, recognized and widespread models or theories will not have comprehensive explanations, as the reader is expected to have a thorough understanding of statistics and financial models. This includes, among others, the assumption of market efficiency and that investors are risk- averse. In addition, the reader is expected to have basic knowledge about ESG.

2 Literature Review

This chapter is devoted to presenting literature that has been considered relevant to the thesis, in order to get deeper insights and knowledge on the subject. First, important concepts that concern responsible investments are introduced, followed by investment screening strategies made by mutual funds and other investors. Next, the importance of investment screening with regard to ESG is described. The literature review is then divided into three distinctions, positive, neutral and negative relationship, for both the theoretical and empirical studies. Finally, the existing literature on controversies and the impact on financial performance are presented.

The literature uses concepts such as SRI, CSR and ESG as synonyms for the same field of study (Waddock and Graves, 1997). Common to all three concepts is that they cover or have an influence on companies' interest in taking on a role as socially responsible actors in society. However, recent studies claim that responsibility and governance can't be analyzed as separate entities, which means that strong governance needs to be included in the concept (Galbreath, 2013; Saltaji, 2013). Likewise, investors are increasingly involving governance criteria into consideration in their SRI analysis. In order to determine outperformance for European portfolios, the governance factor is the most important, while for the North American portfolios, the environmental factor is the most prominent (Eccles and Klimenko, 2019). The SRI concept itself is based on a screening process in which investors seek to be socially responsible by focusing on five different non-financial factors; religion, ethics, environment, society and governance. The five factors are also abbreviated in the literature as ESG, which describes the ethical filter that constitutes a key element to the investment policy (Waddock and Graves, 1997).

In this thesis, ESG investing is defined as an investment strategy that where investors consider extra-financial data. The ESG indicators are used to locate supplementary information of financial performance, that is not directly illustrated in financial figures (Bassen and Kovacs, 2008). This thesis will use the same definition as Bassen and Kovacs (2008): "*The concept of ESG issues refers to extra-financial material information about the challenges and performance of a company on these matters. It thus delivers additional relevant information, allowing more differentiated investment judgements by enabling investors to better assess risks and opportunities*". ESG indicators have become an important tool for measuring non-financial performance of companies, but also to help indicate competencies in a firm's management and assist risk management (Galbreath, 2013).

2.1 Investment Screening

The previous section introduced the five non-financial factors that mutual funds and investors use to target specific investment preferences. As described in section 2.4 on utility functions, the investment decision depends not only on financial considerations but also on the impact of investments on society. This process distinguishes between positive and negative screening strategies that limit the investor to invest within a defined part of the market (Renneboog, Ter Horst and Zhang, 2008).

The 2030 agenda for sustainable development was adopted by all 193 Member States of the United Nations (UN) in 2015, and provides a shared plan for the protection of the planet and improving lives worldwide. This agenda contains the 17 Sustainable Development Goals (SDGs), which stress the importance of responsible actions within a global partnership. The SDGs are meant to provide strategies to reduce poverty, emissions and corruption, while strengthening education, economic growth and overall health conditions (United Nations, 2020).

Figure 2: United Nations 17 Sustainable Development Goals

There are at least two prominent reasons to invest in the SDGs. The first one relates to the return to society. The financial sector has an important social function, namely steering funds toward the most productive investments, after taking full account of social costs and benefits. The second is financial value for shareholders. Companies that provide solutions and contributions to SDGs tend to be better positioned for future changes in the competitive market as well as in the regulatory environments, with lower chances of costly disruption (Schramade, 2017). Most companies are starting to explore

the SDGs and implementing them as a part of their financial reports. The SDGs are increasingly being referred to in companies' communication, but the corporate use of targets and Key Performance Indicators (KPI) are still rare. These kinds of KPIs are needed, as investors are beginning to demand that companies start reporting on their progress on achieving their goals. However, not all SDGs are equally investable and reporting on SDG KPIs is still too scarce to rely on for investment purposes (Schramade, 2017).

From an international perspective, the United Nations has in recent years focused on ESG filters and their influence on investment policy. In doing so, they have designed the United Nations Principles (UNPRI), which are six voluntary guiding principles for responsible investing. The UNPRIs illustrate strategies for implementing ESG initiatives and turning them into actions. Together with the natural consequences of a lack of focus on ESG, the principles help increase the pressure on companies to act responsibly. The development means that a lack of focus on the ESG filters from the companies may potentially have an impact on operations, which will ultimately affect the investor (United Nations, 2020).

United Nations 6 Principles
1. We will incorporate ESG issues into investment analysis and decision-making processes
2. We will be active owners and incorporate ESG issues into our policies and practices
3. We will seek appropriate disclosure on ESG issues by the entities in which we invest
4. We will promote acceptance and implementation of the principles within the investment industry
5. We will work together to enhance our effectiveness in implementing the Principles
6. We will each report on our activities and progress towards implementing the Principles

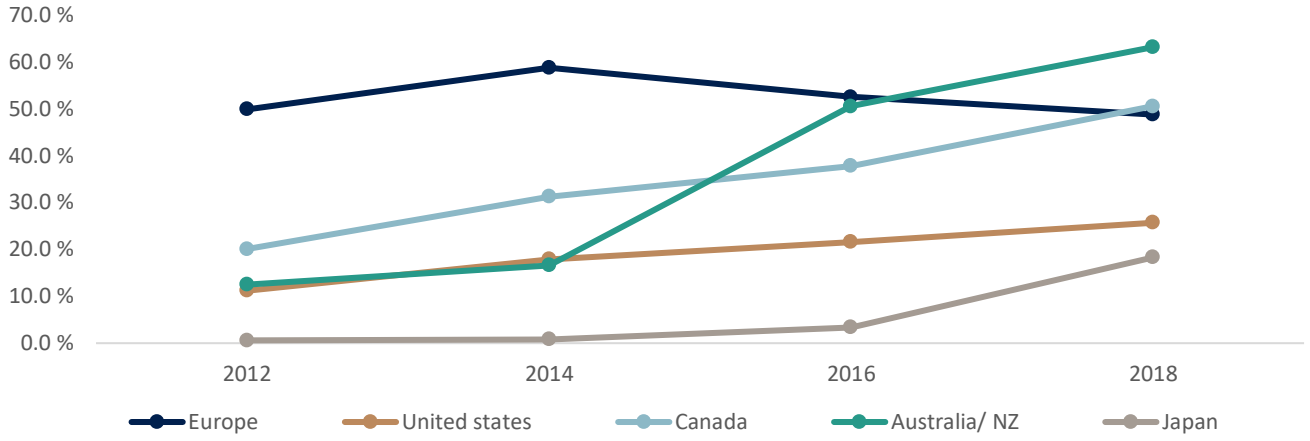
Figure 3: United Nations 6 Principles.

Positive screening implies to select investments in companies that are considered "best-in-class" based on specific environmental, social and governance (ESG) metrics relative to industry peers (United Nations, 2014). Positive screening makes sure that stocks are selected to be included in a portfolio based on the investor's preferences relative to selected parameters. An example of positive screening is when an investor compares two companies with the exact same characteristics and provide the same risk-adjusted return, and chooses to include the company with the highest focus on ESG in the portfolio. The investor is therefore faced with a decision in which the quantitative and

qualitative measures for an attractive investment must be reconciled, and in some cases the investor prioritizes ethics over economics (Bollen, 2007).

Investors' increasing use of positive screening should be seen as a derived effect of the negative screening that was more prevalent in the past (Renneboog, Ter Horst and Zhang, 2008). Negative screening means that investors avoid investing in companies or industries that do little to fulfill the ESG criteria. This means that they have the opportunity to set up corporate criteria for investment decisions, where they do ethical screens and exclude companies that generate revenue from e.g. controversial weapons, nuclear energy, alcohol and tobacco (United Nations, 2014) or exploit employees in developed countries (Renneboog, Ter Horst and Zhang, 2008). These kinds of stocks are frequently described as "sin stocks" (Hong and Kacperczyk, 2009). As many investors restrict themselves from investing in these kinds of stocks, they will often have a higher return requirement and demand a premium. These stocks are less held by norm-constrained institutions and investors, while hedge funds, that are natural arbitrageurs, tend to invest in "sin" stocks (Hong and Kacperczyk, 2009).

Companies are increasingly integrating responsible programs into their business, which in turn yields great attention from mutual funds and investors. Table 1 shows the development of global sustainable investing assets relative to total managed assets from the period 2012 until 2018, and depicts a staggering 34% increase between 2016 and 2018. Furthermore, the figure illustrates that Europe invests more heavily in responsible and sustainable assets than the other four markets, where both the negative and the positive screening contribute to the significant development in the European SRI market.



	2012	2014	2016	2018
Europe	$8 758	$13 608	$12 040	$14 075
United states	$3 740	$6 572	$8 753	$11 995
Canada	$589	$945	$1 086	$1 699
Australia/ NZ	$134	$180	$516	$734
Japan	$40	$53	$474	$2 180
Total	**$13 261**	**$21 358**	**$22 869**	**$30 683**

Table 1: Snapshot of global sustainable investing assets relative to total managed assets, 2012-2018. The line chart depicts the proportion of sustainable investing relative to total managed assets in each respective region, 2012-2018. Numbers in billions. Source: Global Sustainable Investing Review 2018/2016. Authors own illustration. (The report is Biannual and more recent numbers are not yet available).

2.2 The Importance of ESG Screening for Companies

The following sections will lay out relevant incentives as to why companies initiate ESG measures which exceed what is required by law. As discussed, investment screening as a tool is becoming increasingly more common for mutual funds. As the number of investors who desire to invest socially responsible increases, not ignoring these issues are becoming a natural part of companies' self-interest. Companies which ignore socially responsible initiatives may face unwanted pressure from stakeholders, including regulatory authorities, which resultingly could imply increased costs. Moreover, being ahead of future legislations may contribute to a competitive advantage, if the companies simultaneously succeed in differentiating from competing businesses (Rennebog, Ter Horst and Zhang, 2008).

2.2.1 Should Companies be Socially Responsible?

As discussed in the aforementioned, there is an increasing focus on ESG parameters among investors. The development has made them consider the external implications of their investment decisions. The advantage and disadvantages created by a company's actions that affect another agent in the economy are defined as externalities (Renneboog, Ter Horst and Zhang, 2008).

Negative externalities occur as a result of actions where profit maximizing behavior impacts the surroundings and thus society negatively. For example, a company emitting toxins in the local watercourse and thereby destroying the harvest from surrounding fields, will be the source of a negative externality from the company's operations. Renneboog et al. (2008) emphasize that there often is a conflict between the financial merits of an investment and the associated negative externalities. Financial solutions to the obvious problems of externalities have been the development of markets where externalities are tradeable, with the purpose of internalizing them. Among other initiative, trade in CO_2 quotas is an attempt at internalization. Moreover, companies may choose to internalize negative effects by performing investments in CSR and in that way enhance their ESG profile.

The classic utility perspective describes an investor which seeks to maximize his returns by optimizing the relationship between risk and return, without considering qualitative parameters in their investment decisions (Markowitz, 1959). However, as argued in section 2.3, new theories have contributed to alternative utility functions where the investment decision does not solely rely on the classical understanding of the risk-return ratio. In this sense, non-financial parameters, including ESG, are included in the investors' utility function. This means that investors need to consider a trade-off between their financial and qualitative interest, where both profit maximization and benefits of the investments are evaluated.

Renneboog et al. (2008) further argue that an expansion of the utility function could contribute positively. If a responsible investment generates a positive net present value for shareholders while minimizing negative externalities, it could help increase the share price as it maximizes utility in several parameters. However, how to maximize these parameters and how investors prioritize financial returns and responsibility is difficult to answer. This trade-off for companies is essential to the ESG screening criteria. As the criteria influence the investment decision of responsible investors, it is crucial to uphold them to be included in the investment universe, but also provide attractive risk-

adjusted returns. Renneboog et al. (2008) emphasize that the stock market is under-assessing the value of ESG investments in the short term, although studies have found that a proactive ESG strategy may be beneficial. In the longer run, there are significant downsides associated with neglecting social responsibility, both in terms of litigation and reputational risk.

2.2.2 Why do Companies Implement CSR Strategies?

Although there is no consensus among researchers on what role ESG should play in the overall corporate strategy, ESG strategies have become a regular element of most companies of a certain size. The exact motivation for implementing such strategies is unclear. Existing literature has suggested that initiatives may emerge as a result of agent costs. Moreover, it is reasonable to assume that stakeholder pressure varies depending on company size, with the larger ones experiencing heavier scrutiny. A common way to employ an ESG strategy, is to enact in CSR.

A popular argument for initiating a voluntary CSR policy is its direct connection to maximization of shareholder value. Heal (2005) opines that CSR minimizes conflicts between the business and the surrounding society, and thus reduces agent costs. Moreover, Allen et al. (2007) discuss shareholder oriented companies versus stakeholder oriented ones, and find that the latter on average have higher valuations.

Several studies claim that CSR might help minimize asymmetric information in both financial and labor markets. Having a strong CSR brand may be used to signal a solid workplace, which can help attract motivated and skilled employees (Brekke and Nyborg, 2004). Companies may use CSR to make it easier for investors to judge their reputation (Brekke and Nyborg, 2004). CSR is then used as a tool to signal that the company is trustworthy and offers quality products (Fisman, Heal and Nair, 2006).

Contrary to the arguments that companies voluntarily implement CSR strategies because it increases the value of the company, theories have argued that CSR strategies are more a result of pressure from the surrounding community and environmental lobbyists. Baron (2001) describes lobbyists pushing companies to adopt a more rigorous approach to environmental standards. Naturally, larger shareholders have greater influence and may use their position to act as social activists, as they themselves often are under scrutiny from the outside world. Likewise, the regulatory environment

and other societal actors are increasingly influencing business decisions, which may be the explanation for an increased focus on ESG in general and CSR policies more specific.

Nevertheless, stakeholder pressure is a yet more nuanced picture. Darnall et al. (2010) emphasize that SMEs respond differently to the same type of pressure. Smaller businesses have a weaker response to outside pressure, which, according to Darnall et al. (2010), is explained by them having relatively fewer resources and a more simplified decision making process than larger companies

Bianchi and Noci (1998) support this notion and suggest that a proactive CSR strategy is problematic for SMEs as they lack the resources and skills to implement them. It will often be difficult for management to justify such investments rather than investments in core business operations. This is because these types of investments do not necessarily generate cash flows for the business in the short or medium term (Bianchi and Noci, 1998).

The argument has also been made that CSR is the result of conflict of interests. Contrary to the above, Barnea and Rubin (2010) argue that CSR is the result of agent costs between management and shareholders. Barnea and Rubin (2010) highlight that the management may over-invest in CSR for their private benefit, specifically for the purpose of improving their reputation as good and responsible citizens. It maximizes their individual benefit as the management may desire being associated with companies that are socially responsible. The conflict may be limited if the management also are shareholders, and thereby bear part of the cost of value eroding initiatives. Moreover, Barnea and Rubin (2010) argue that value eroding initiatives can be limited by lowering free cash flow. Agent costs can thus be reduced simply by increasing debt, because the free cash flow that management dispose of decreases. In this regards, increased debt has a disciplinary effect on management.

In summary, the literature has identified three main reasons why companies implement CSR strategies:

Incentives for implementing CSR strategies
• Companies can voluntarily choose to implement CSR strategies to lower the competitive intensity in products markets and increase company value. • CSR strategies can be used to signal quality and credibility, thus addressing both product buyers and also help attract motivated and skilled employees. • Involuntary CSR strategies can be implemented due to pressure from lobbyists or as a result of agent costs, where management over invests for their own gain. Given that companies' sole purpose is to maximize shareholder returns, it is important to identify whether investors are willing to pay for CSR, as a company should only invest in social responsibility if it can help maximize shareholder returns.

In the next sections the authors will elaborate on different research that have been investigated, and divide them into positive, neutral and negative relationship in regard to ESG and financial performance.

2.3 Theoretical Perspective on ESG Investments

Bollen (2007) argues that investors have a multi-attribute utility function, which includes additional elements to the relationship between risk and expected return, when considering possible investments. The paper defines the extended utility function as a portfolio's return distribution, as well as a variable that represents whether it is a socially responsible investment decision. These attributes could be SDGs that a fund is working towards achieving.

This assumption is supported by Statman (2000), who argues that the existing literature on behavioral finance considers the investment decision as a type of product choice, where "*value-expressive*" characteristics of the product may influence its attractiveness. This perception stands in stark contrast to the standard paradigm of financial theory, where all investment decisions are incentivized by optimizing the relationship between variance and return.

Renneboog et al. (2008) opine that the relationship could be explained by capital markets internalizing firms' externalities, cf. paragraph 2.2.1, and investors are willing to pay for corporate social performance. However, if internalization does not have an effect, the explanation could be that environmental initiatives by the firm is correlated with future cash flows (Renneboog, Ter Horst and Zhang, 2008).

2.3.1 Theoretical Positive Relationship

Multiple theories argue for a positive relationship between a firm's investments in and prioritization of ESG related issues and financial performance, under which the resource-based view is the most prominent. The resource-based view states that sustained competitive advantage may be derived from investing in environmental performance and social responsibility. Divestments of particularly polluting business areas before they become targets for potential lawsuits could result in better performance for the firm, due to the management proactively addressing the negative externalities of their business strategy (Russo and Fouts, 1997; Schmidt et al., 2015). Further, in light of the resource based view, Porter (1995) argues that an active environmental policy may contribute to competitive advantage as the firm is able to achieve a more cost efficient use of its resources. The reasoning behind the argument is that the financial advantages the firm achieves by making strategic environmental investments, may be used to generate an improved input-output efficiency and develop new business areas (Derwall et al., 2005). Similarly Additionally, Barney and Hansen (1994) argue that firm specific resources, which are non-substitutable, may provide the company with sustained competitive advantage, which again may contribute to financial advantages. Consequently, the rationale proposes that firms which value ESG initiatives are able to achieve higher returns than firms that do not (Derwall et al., 2005). The theoretical foundation for this theory is that environmental and social responsibility is a performance dimension on which stakeholders evaluate the firms (Schmidt et al., 2015).

Further building on Schmidt et al. (2015) and Porter's (1995) argument that environmentally responsible investments may contribute to sustained competitive advantage, the access to new markets is a central consideration.

Firms that invest in improving their ESG profile may gain access to capital markets which their less responsible equivalents will not. The demand for "greener" products is steadily increasing, which

allow firms to use differentiation strategies to exploit market segments where consumers are getting more aware of their environmental impact (Ambec and Lanoie, 2007).

Furthermore, Ambec and Lanoie (2007) argues that responsible firms may obtain cost reductions by:

- Risk management and stakeholder management
- Material, energy and service costs
- Cost of capital
- Labor costs

Accordingly, socially and environmentally responsible investments may result in both cost and revenue advantages for the firm.

The literature further argues that green investments imply good overall management, future value creation and innovation (Climent and Soriano, 2011). This suggests that a firm's environmentally responsible investments may be used in screening processes of firms' operations and risk, which represent valuable information to potential investors.

The relationship between firms and their relevant stakeholders is an area of increasing interest in both the academic and business world. Neo classical economists generally argue for a positive correlation between environmental and financial performance (Freeman and Reed, 1983; Barney and Hansen, 1994; Hill, 1995; Berman et al., 1999). The relationship is supported by an argument that environmental considerations can contribute to reduction of costs to external interests and NGOs. A reduction of conflicts with these interests could imply higher returns and make environmentally responsible firms more attractive to investors.

Freeman (1983) is a central contributor to the theory, and his findings conclude that a company should operate within the boundaries of a specific environment, where strategic decision-making is based on interests of all stakeholders and not solely on economic concerns. Through methodical prioritization of various stakeholder interests, as opposed to taking a passive position, a company may achieve economic benefits (Freeman and Reed, 1983). Similarly, Barney and Hansen (1994) Barney et al (1994), Hill (1995) and Berman et al. (1999) argue that solid relationships to the company's stakeholders may reduce associated costs significantly and hence positively affect the financial performance.

Furthermore, green investments may reduce the risk associated with future regulatory changes. Companies can benefit from this by aspiring for higher standards and leverage their first mover position in the market. Additionally, it could arguably be easier to obtain the necessary approvals to invest in new production facilities and machinery with a healthier relation to local authorities (Ziegler, Busch and Hoffmann, 2011). The latter point is especially important for companies operating in countries with a higher degree of government influence.

2.3.2 Theoretical Neutral Relationship

Classical theories such as the asset pricing theory and the efficient market hypothesis prescribe that return is proportional to the level of associated risk. Hence, it should not be possible, through ESG screening, to obtain significant excess returns on responsible assets relative to companies with equal risk profile but inferior ESG scores. Hoepner (2010) challenges the consensus view of existing literature that an ESG filter, everything else equal, leads to a poorer portfolio diversification and a lower risk- adjusted return. Hoepner (2010) further emphasizes that diversification is dependent on three factors;

- Number of selected assets
- Correlation between selected assets
- Standard deviation of selected assets

Hoepner (2010) further argues that a portfolio manager seeking to minimize risk by diversifying his portfolio using stocks with low ESG scores, in reality achieves the opposite, as lower ESG scores are associated with higher risk. This argument is supported by Schröder (2007) and Bello (2005) , who study investment fund portfolio diversification and the standard deviations of SRI and conventional indices respectively. The effect is offset by lower diversification as companies that have a high ESG performance have lower unsystematic risk than companies with low ESG performance.

McWilliams and Siegel (2000) criticized several studies for misspecification in measuring CSR's impact on financial performance. They emphasize that previous studies regressed corporate ESG performance and other control variables concerning financial performance without making, according to them, appropriate adjustments. McWilliams and Siegel (2000) argue that investments in R&D activities should be adjusted for, as it has been shown to have significant impact on firms' financial performance. Studies which do not adjust for this effect, will have an upward bias in their estimates.

When McWilliams and Siegel (2000) adjusts for R&D investments, the results show that CSR has a neutral impact on financial performance.

2.3.3 Theoretical Negative Relationship

That responsible investments impose higher cost of capital to companies, is a well-known and established argument. From a traditional neo classical perspective, it is argued that environmental and social consideration may lead to additional costs which in turn affects financial performance negatively (Oberndorfer et al., 2013). This is partly due to the fact that companies' cost of such investments exceeds the estimated economic benefits. Companies, which do not solely focus on maximizing shareholder wealth, and invests in environmental efforts or make donations to environmental organizations, should experience lower relative returns (Becchetti and Ciciretti, 2009). However, only a smaller portion of the total investor base prefers to invest in responsible companies. As traditional investors do not have this preference, stock prices will likely be negatively affected if a company decides to make ESG friendly investments, as they are considered to incur an unnecessary additional cost.

However, there are still incentives for the company's management to perform responsible investments. This is due to the fact that implicit agency problems between the management and shareholders often occur. Ziegler, Busch and Hoffmann (2011) opine that only if the appropriate management and incentive systems are implemented in the company structure, will the management choose to maximize profits for shareholders. Accordingly, under certain circumstances, the shareholders will endure a situation with lower returns on their investment, as the management is rather optimizing their own utility. This could be investments in environmental projects which do not benefit the shareholders positively (Ziegler, Busch and Hoffman, 2011). Due to these concerns, Ziegler, Busch and Hoffmann (2011) argue that investors could benefit from avoiding companies which invest in responsible projects, but rather have a purely profit maximizing approach.

Yet another theoretical argument for a negative relationship is derived from the theory of diversification. A frequently highlighted problem in making investments in companies with a high degree of environmental considerations is that such strategies make the relevant investment universe limited and thus reduces the ability of the investor to diversify their portfolio. Green and SRI funds should consequently be associated with higher risk as their mandate put restrictions on the relevant

investment universe and reduce the possibility of diversification (Renneboog, Ter Horst and Zhang, 2008).

As a result of these investment restrictions, responsible funds are developing portfolios with an overweight in equities with unattractive risk-adjusted returns to meet predefined ESG standards. This theoretical argument may be considered as a possible explanation for some studies finding that green funds show lower returns than traditional funds (Climent and Soriano, 2011). However, the theoretical argument is criticized by, among others, Hoepner (2010), who described that lower unsystematic risk for companies with high environmental performance offsets the negative effects of diversification.

2.4 Empirical Studies on ESG

This section is devoted to presenting existing literature about the effect of ESG and its impact on business specific measures. Studies regarding ESG will be looked upon, in addition to studies that concern ESG controversies.

Several studies have used theoretical discussions about the relationship between a company's responsible and financial performance in order to analyze an empirical relationship. The studies that have been conducted have different approaches when analyzing the respective problems, hence finding deviating results. Based on the previous theoretical arguments, the different study designs and the differences in their results will be explained. According to the review of the theoretical arguments for positive, neutral and negative correlations between ESG and financial performance, the empirical data is divided in relation to the respective studies' conclusions.

2.4.1 Empirical Positive Relationship

Zhang (2017) published a literature study that was formulated as a recommendation to board members of Dutch pension funds. She argues that responsible investments do not restrain traditional investments, and that responsible investments can generate excess return by utilizing mispricing of ESG factors. Furthermore, she explains that companies with high ESG scores have outperformed their benchmarks. This is true for elements which cover *corporate governance*, *ecoefficiency* and *employee relation*. On a company level she refers to studies that show that companies with high ESG score have a stronger financial performance in addition to lower cost of capital. However, she

concluded that these factors could already be priced in, as a result of the increased focus on ESG (Zhang, 2017).

Similar to Zhang (2017), the study made by Harjoto and Jo (2015) find that a company's share price volatility and cost of capital decreases when their ESG activities increase. Sharfman and Fernando (2008) find that companies who tends to have a strategy based on environmental risk management are rewarded with a lower cost of capital. El Ghoul et al. (2011) find that firms with higher ESG score achieve significantly lower cost of equity capital, thus indicating that companies with a higher ESG score will have lower risk. Gregory, Tharyan and Whittaker (2014) confirm that companies with high ESG scores enjoys a lower cost of equity capital, both on the short and long term. They find that companies that classify as "green", i.e. companies that invest heavily in environmental initiatives, have a higher long-term growth potential compared to companies that are "toxic", i.e. companies that have many concerns and damage the surroundings.

Chollet and Sandwidi (2018) investigate the complex relationship between the financial risk of firms and their engagement in ESG, measured by Thomson Reuters Asset4 scores. They analyze an international sample of 23.194 firm-year observations on 3.787 worldwide firms, with observations from 2003-2012. The results show that a firm's good social and governance performance significantly reduces its financial risk, and thereby reinforces its commitment to good government and environmental practices. Similar to Ioannou and Serafeim (2012), which notice a strong correlation (0,77) between environmental and social scores, which suggests multicollinearity among these variables. In general, ESG performance scores are negatively and strongly correlated with firm risk (beta, specific risk, total risk), which is consistent with the findings from Bouslah et al. (2013).

Russo and Fouts (1997) analyses 243 firms and find that there is a positive correlation between environmental performance and return on assets (ROA). Furthermore, they find that the correlation is stronger in high-growth industries where the companies also are expected to have stronger compliance measures and policies, i.e. stronger governance. Hart and Ahuja (1996) find that there is a significant positive correlation between the company's bottom line and its pollution reduction initiatives. More precisely, they investigate the impact that environment friendly initiatives have on

the company's operating ratios such as *Return on Sales (ROS)*[4]*, Return on Assets (ROA)*[5] and *Return on Equity (ROE)*[6]. The study explains that companies with high pollution are the ones benefitting the most from environmental improvement initiatives, because these companies can more easily reduce their environmental damage. In other words, the better a company performs environmentally, the more expensive it will be to implement further environmental improvement initiatives. Hart and Ahuja (1996) argues that the marginal costs of investing in environmental improvement initiatives rarely outweigh the marginal positive effects of these investments. The companies that have significantly lowered their environmental pollution will discover that the positive effects of investments will outweigh the costs. However, the data and the results are limited by the fact that the period for the analysis is from 1988-1989. At this point in time, only a small amount of pollution-heavy industries, such as shipping and transport, had undertaken initiatives for environmental improvement (Hart and Ahuja, 1996).

Friede, Busch and Bassen (2015) combines over 3,700 study results were approximately 2,200 are unique primary studies. Their research finds that almost 90% of the studies reports non-negative findings between ESG and corporate financial performance (CFP). They use a vote- count study, simply comparing the number of positive, negative and neutral research of studies, for non-portfolio- and portfolio-studies to measure the relationship between ESG and CFP. For the non-portfolio primary studies, the results finds 56.7% positive, 5.8% negative and 18.8% neutral correlation between ESG and CFP. However, when they analyze the studies for primary portfolio studies, they find a different allocation. These studies finds 15.5% positive, 11% negative and 36.1% neutral correlation between ESG and CFP. Their meta-analysis finds that the large majority of studies shows positive findings and were particularly evident in North America (42.7%) and in emerging markets (65.4%). This relationship seems to have been stable since the early 1990s. When looking closer into developed markets in Europe, only 26.1% of the studies show a positive correlation, while 8% show a negative correlation. Friede, Busch and Bassen (2015) point out that portfolio studies tend to give

[4] ROS= Operating profit/ sales

[5] ROA= Net income/ average total assets

[6] ROE= Net income/average shareholders' equity

neutral results because of the inclusion of management costs, composition restrictions and that portfolios are subject to both systematic and unsystematic risk. Revelli and Viviani (2015) agreed with the latter and pointed out that ESG screening involves increased costs. It is possible that these conditions will distort the effects of ESG initiatives in portfolio studies. This means that if the benefit from responsible investments is small, the investor will most likely have trouble of retaining it after implementation costs.

During the last 20 years, several ESG initiatives have emerged, and one of the most prominent and widespread initiative is the United Nations Global Compact (UNGC). Ortas et al. (2015) investigates the impact of UNGC commitment on corporations' ESG performance, and furthermore analyses the relationship between the corporations' ESG performance and their financial performance. The study is based on panels of data which cover companies that are operating in Spain, France and Japan, as these countries have the highest number of UNGC participants. Their empirical analysis shows that the companies that stay committed to the UNGC experience a positive and significant impact on their ESG performance. However, it is worth mentioning that the observed period is from 2008-2013, where the global economy entered the deep recession, and thus indicates that the positive implications are present even in bear market periods. Another interesting finding is that the companies committed to UNGC saw a significant and positive relationship between their ESG performance and their financial performance, which in this study was measured by the companies return on investment (ROA) and Tobins Q (TQ) (Ortas et al., 2015).

2.4.2 Empirical Neutral Relationship

Cohen et al. (1997) find evidence that can answer the question concerning that "green investors" would need to pay a premium for their investments in sustainable companies. They find that if an investor combines only the environmental leaders in a portfolio that is focused on one specific industry, it will yield the same return, or better, as if the investor would combine environmental laggards from different industries in their portfolio. They investigated the relationship between environmental- and financial performance by combining two sets of portfolios. In their analysis, they formed a "low-pollution" and "high-pollution" portfolio using S&P500 and IRRC Legacy Governance data, where they measured on eight environmental parameters. Their study formed six company-size and industry balanced portfolios in the two categories and tested the financial

performance in terms of ROA and ROE, together with risk-adjusted return on both of them. Their results show that when measuring risk-adjusted ROA and ROE, the "low-pollution" portfolios have a better performance compared to the "high-pollution" portfolios, which was shown in over 75% of the portfolio comparisons. However, only 20% of the results produced were statistically significant. Based on these results, there are no strong evidence of a green investment premium. Instead the findings can be interpreted as evidence that investors do not forgo excess return by investing green, and that investors who build balanced portfolios based on companies with strong environmental performance will experience any lower return on investment compared to benchmark.

Halbritter and Dorfleitner. (2015) investigate the link between corporate social and financial performance based on ESG ratings, and review the existing empirical evidence for this relationship. The ESG data is collected from Bloomberg, KLD and Asset4 for the US market in the period 1991-2012. They combine companies with different ESG scores in portfolios and perform their analysis by applying the four-factor model and Fama and MacBeth regression. The study tests the robustness of the results with both market capitalization weighted and equally weighted implementation. The results do not provide significant differences in returns for the ESG portfolios that were divided into companies with high and low ESG scores, respectively. However, the Fama and MacBeth regressions revealed that several of the ESG variables influenced the result, but the influence is strongly dependent on the particular ESG rating company. They therefore do not identify an unambiguously significant impact on the financial performance of the three ESG score providers, neither on the overall scores nor on the specific environmental scores. Their study suggests that investors should not expect abnormal returns by trading with high and low rated firms in regard to ESG (Halbritter and Dorfleitner, 2015).

Galema et al. (2008) investigates the effect of ESG on stock returns by using six different ESG parameters from KLD Research and Analytics in the period 1992-2006. The study tests if any of these portfolios can deliver excess return by using Fama & French three- factor and the four-factor model. Their initial results confirm, similar to previous research, that the risk-adjusted over performance of ESG stocks is not significantly different from zero. Next, they find that some ESG scores influence the book-to-market ratio. The authors find that the aggregate results of the analysis of SRI scores could eliminate a relationship if, for instance, some of the individual dimensions of SRI have opposing effects on the financial performance. A common approach when analyzing stock returns and controlling for risk, is to use Fama & French regression which include a high minus low (HML)

factor. This factor measures the sensitivity of the differences on stock returns for stocks with high and low book-to-market ratios. They conclude that socially responsible investments for companies result in lower book-to-market ratios, where the SRI effects do not influence the alphas. The study does not find statistically different performance for companies that engage in SRI compared to conventional ones (Galeme et al., 2008).

2.4.3 Empirical Negative Relationship

There are few studies that find a significant negative relationship between responsible initiatives and financial performance, compared to the studies that find a positive or neutral correlation. Nevertheless, these studies are important to analyze, as the theory addresses a negative relationship between the two (see section 2.3.3). In this section, the authors will present empirical studies that support the theoretical arguments of a negative context.

Oberndorfer et al. (2013) analyze companies that is included in sustainability stock indexes and their stock performance. The study is based on Fama & French three-factor model, in addition to a t-GARCH (1,1). Their results suggest that stock markets may penalize the inclusion of a firm in sustainability indexes, where their findings are mainly driven by a strong negative effect of the inclusion in the Dow Jones Sustainability World Index. Furthermore, the study implies negative average cumulative abnormal returns, where it will have larger negative impacts by being included in a more visible and well-known sustainability stock index.

Bauer, Koedijk and Otten. (2005) review and extend previous research on ethical mutual fund performance. They apply a standard CAPM and a four-factor model which controls for size, book-to-market and stock price momentum, in order to overcome the benchmark problem that most prior ethical studies have suffered from. Their study reports three interesting results. Firstly, it is no proof of a statistically significant difference in return between ethical mutual fund and conventional fund returns, after completing adjustments for factors such as book-to-market, size and momentum factors. Furthermore, ethical mutual funds seem to have dissimilar investment preferences compared to conventional funds. Ethical funds also seem to be less exposed to market return variability compared to conventional funds. In the UK and Germany, ethical funds are in a larger extent exposed to small capitalization firms, while ethical funds in the US are investing more heavily in large capitalization firms. This implies that ethical mutual funds tend to have a higher focus on growth companies in

Europe, while it is a more value-oriented perspective in the US. Lastly, they find that standard indices are a better tool for explaining ethical mutual fund returns compared to ethical indices (Bauer et al., 2005).

Berman et. al (1999) studies the relationship between stakeholder management and firm financial performance. The study was conducted using KLD database, one of the first socially responsible investing indexes, which uses five variables in order to represent a firm's stakeholder posture. These variables are *employees*, *product safety/ quality*, *community*, *diversity* and *natural environment*. Only employees and product safety/ quality directly affected the financial performance, while the three others had no significant impact. These results reinforce the perception of stakeholder theorists that emphasizes the importance of a firm's relationship with its employees and customers. In addition, these two variables may be used as a source of differentiation for an individual firm in order to achieve a higher financial performance. On the other hand, the study did not find any significant results for community, diversity and natural environment in connection to financial performance. According to the authors, a plausible explanation for this may be that their sample included firms from many different industries, hence it is likely that regulations for environmental initiatives do not have uniform impact across industries. The study concludes with that there is a negative relationship between better environmental performance and the company's financial performance (Berman, Wicks, Kotha and Jones, 1999).

Aupperle, Carroll and Hatfield (1985) studies the relationship between CSR and profitability. They found that there is an inverse relationship between responsibility and financial performance, where the emphasis on one of the two components was on the expense of the other. Apparently, the more concerned a company was with its financial and economic affairs, the less focus was directed at responsible actions. Waddock and Graves (1997) argue that the rationale behind this result is that the companies increase their capital expenditure on initiatives that optimally should be paid by other entities. This could be, for instance, that a company chooses to invest in equipment that reduces the overall emission, when other entities, such as competitors and the government, resist to incur these costs that will benefit the entire society. According to Friedman (1970), there are numerous costs for responsible behavior, while there are only minor economic benefits. By this argument, the costs of responsible initiatives drop directly to the bottom line of a company's financial statement, simply reducing profits and thus reduce shareholder wealth.

2.5 Controversies

An ESG score comprises multiple measures, and controversies related to ESG is often included (Thomson Reuters, 2020). To extract controversies and study the measure through a separate score, is a newer method that is not covered by existing literature to the same extent as conventional ESG scores have been. Thomson Reuters first started measuring controversies separately when they replaced their ASSET4 scoring system with Refinitiv in 2018. The literature review indicates that relatively few studies have been conducted on the use of controversy as a separate score or assessment criterion. The key findings of previous studies are summarized in the following.

Aouadi and Marsat (2018) argue that ESG and controversies related to ESG are not direct contradictions. They examined the effect of controversies in a study of over 4000 companies from 58 countries. The study is based on regressions with Tobin's Q as a dependent variable, including controversy and ESG scores as independent variables. They find that being involved in multiple controversies is associated with a significant excess return when looking at "high-attention" companies. These companies are characterized by being bigger, doing better, getting more attention from investors and located in countries with a higher degree of freedom of the press. This indicates a positive relationship between controversies published about a company in the media and Tobin's Q. This positive relationship is arguably not intuitive. However, the direct effect of the controversies is not significant when looking at ESG scores and controversies together. Despite the negative publicity in the media, the authors claim that the visibility of the company will increase, and thus also the visibility of their ESG score. This, they claim, will reinforce the effect the ESG score has on the company's value. For "low-attention" companies, the effect of controversies is not significant, either directly or in combination with ESG scores.

Another interesting study was conducted by NN Investment Partners and the European Center for Corporate Engagement at the University of Maastricht in 2016. Their findings concluded that by excluding companies that have been involved in controversies, a higher Sharpe ratio was obtainable, i.e. higher risk-adjusted return (NN Investment Partners, 2016).

A study conducted by Blackrock (2018) found that the companies that reported having the most guidelines related to ESG were also involved in the most controversies. On the other hand, those who reported relatively few ESG policies were subject to fewer controversies. Future controversy was also positively correlated with the size of the company and how many existing controversies they were

already involved in. Further, as a response, controversies often lead to the implementation of additional policies related to ESG, but usually no noticeable change in practice was recorded. In a previous study on behalf of Blackrock, Garvey et al. (2016) argue that controversies have a significant negative effect on returns. This effect becomes both greater and more significant when controlling for size, industry and country. They therefore argue that controversies can be used as an instrument of reputational and regulatory risk.

2.6 Key Takeaways

Literature Review

- Existing literature does not offer a consensus view on the relationship between ESG and financial performance. Results are especially dependent on ESG data provider, geography, time period and used market portfolio.
- The literature presents various rationales for how ESG may contribute to returns, whereas the reduction of risk is the most prominent.
- ESG considerations may contribute to an expansion of the investor utility function, where other elements than the optimization of risk and return is considered and desired.
- ESG performance is explicitly linked to company size and emphasize on the relative overperformance of larger companies. This is explained by them having more resources to employ ESG initiatives which do not necessarily generate positive cash flows in the short and medium term.
- Likewise, controversies are also linked to size, with larger companies being more exposed. Studies explain this relationship by arguing that larger companies experience more thorough scrutiny from stakeholders.

3 Model Specification

This chapter will elaborate on the motivation for the models used in the analysis. Next, each individual factor model will be explained in detail. The final section will discuss the idea of portfolio diversification, with respect to systematic and unsystematic risk.

3.1 Motivation for used Models

Previous studies have employed various methodological approaches and models to test the long-term relationship between ESG and financial performance. Alternative models, such as t-GARCH (Becchetti and Ciciretti, 2009), (Oberndorfer et al.,, 2013), and a simple correlation matrix (Waddock and Graves, 1997) are examples from previous research. However, the literature that have investigated the presented subject, has generally used three specific models. These includes the CAPM and Carhart four-factor (Ziegler et al., 2011), (Bauer et al., 2004), (Renneboog et al., 2008), in addition to the Fama & French three- factor (Derwall et al., 2005), (Halbritter and Dorfleitner, 2015). The meta-study by Revelli and Viviani (2015) also includes the use of the Fama & French five-factor model.

The authors of this thesis find that both the three- and four-factor model have been widely used in order to increase the explanatory power of the results relative to the results from the CAPM. This approach stems from the desire to treat the series of anomalies that have been addressed and identified by Fama and French (1992). The literature that investigates the problem area for this thesis is still limited and shows deviating results. There is a lack of supporting and comparable empirical studies on the subject, and whether the models produce a satisfactory degree of explanatory power. This thesis will attempt to apply the CAPM, Fama & French three-factor, Carhart four-factor and Fama & French five-factor. By using these models, the authors believe that this study can contribute to a more nuanced analysis of the relationship between ESG and financial performance and support the existing empirical results. In addition, these models will be used to test controversies and its impact on financial performance, which is a fairly new area of research. The following chapters will give a detailed description of each individual factor model that is employed.

3.2 CAPM

Modern portfolio theory was developed by Markowitz (1952, 1959), where several authors in the 1960s, including Treynor (1962), Sharpe (1964), Lintner (1965) and Mossin (1966), have been part

of reshaping the theory of how portfolio managers approach portfolio risk. In essence, modern portfolio theory starts off with a suggestion that portfolio risk is determined by the co-variances of assets that are included in the portfolio. By the mid-1960s, the modern portfolio theory evolved to include the Capital Asset Pricing Model (Zhang, 2017). The aim of the model is to describe the relationship between systematic risk and expected return on assets, assuming that the return on investment portfolios is proportional to the associated risk and that an optimal portfolio is well diversified (Derwall et.al., 2005). It is based on the principle that all investors have to be compensated for the time value of money and the risk involved when buying stocks. The model consists of the risk-free rate, beta and the market risk premium. In practice, return on Treasury bonds are used as this is the closest you will get to a risk-free placement of assets. The compensation for the systematic risk is given by the beta of the market multiplied by the market risk premium, i.e. excess return beyond the risk-free investment. If a stock is exposed to high systematic risk hence have a high beta, you would demand a compensation such as higher expected return (Zhang, 2017). The model is expressed by the following equation:

Equation 1: CAPM

$$E\ (R_i) = \ R_f + \ \beta_{iM}\ (E(R_M) - R_f f)$$

CAPM relies on a market-related measure of risk, namely the market beta, in order to price a portfolio of assets. Empirically, beta for asset *i* can be expressed as:

Equation 2: Market beta

$$\beta_i = \left(\frac{\sigma_i}{\sigma_M}\right)(r_{i,M}) = \frac{Cov\ (R_i, R_M)}{\sigma_M^2}$$

where $Cov\ (R_i, R_M)$ is the covariance between the return on asset i and the return on the market portfolio. The CAPM model assumes that the market portfolio is efficient in a return-variance context. The implication of this assumption is that the differences in expected returns across the shares can be fully explained by the differences in the beta measure. This means that no variables other than beta should have any influence of the explanation of the expected return. The CAPM is based on a multitude of underlying assumptions, including the assumption of an efficient market, which was presented by Fama (1965, 1970, 1976) as the efficient market hypothesis, a theory that have been thoroughly debated in the literature (Heymans and Brewer, 2015). Some of the criticism concerns

how the risk-free rate is used in the model. The issue is that the yield changes daily, i.e. creates volatility, and the assumption that shareholders can borrow and lend at a risk-free rate is not possible in practice. Also, the CAPM predicts that asset risk premiums only depend on the asset's beta and that the only factor that matters is the market portfolio, which gives a doubtful view of the risk involved (Ang, 2014). Some of these concerns have been adjusted through the use of Jensen's alpha and through the work from Fama & French three-, and five factor models, which will be discussed in the following subsections.

3.2.1 Jensen's Alpha

As a further development of the CAPM model, Michael C. Jensen (1968) introduced a risk-adjusted performance measure (alpha) that represents the average return on a portfolio or investment that will be above or below the estimate from the CAPM. Alpha will determine the abnormal return over the theoretical expected return, that is not explained by the expected market return. The error term will then be expected to be zero. Jensen's alpha is expressed in the following formula:

Equation 3: Jensen's Alpha

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

A strategy based on buying and holding the market will result in an alpha equal to zero. An investor who manage to choose stocks that on average beats the market, will have a positive alpha. However, if an investor chooses stocks that performs worse than the market, it will result in a negative alpha (Jensen, 1968).

3.2.2 Dependent Variable

The dependent variable in the model is a continuous numeric variable and is stated as the excess return on the portfolio ($r_{i,t} - r_{f,t}$). $r_{i,t}$ is the monthly realized return on the value-weighted portfolio, where the average realized return on the risk-free asset is deducted. In the CAPM, it is assumed that the realized value of alpha will be equal to zero for the sample of historical observed returns. The dependent variable is applied in the following factor models.

3.2.3 Independent Variables

The two explanatory variables in the model are used in order to explain the dependent variable and are estimated in the regression analysis.

- **Alpha, α_i,** is the average return on the portfolio above (or below) what is predicted by the CAPM model, after the portfolio's beta and the average excess return on the market is accounted for.
- **Market beta, β_1,** is the market risk premium for exposure to systematic risk

3.3 Fama & French Three-factor Model

Fama and French (1992) developed an asset pricing model that expands the CAPM by adding size risk and value risk factors. They established three stock market factors: an overall market factor and factors that relates to book-to-market equity and firm size. Throughout their research, they found that value stocks have a tendency to outperform growth stocks, and that small-capitalization stocks outperform large-capitalization stocks. The three-factor model adjusts downward for small-cap stocks; hence the performance of a small-cap portfolio will tend to have a lower result compared to the CAPM.

The three-factor model is expressed through the following equation:

Equation 4: Fama & French three-factor

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_1(r_{m,t} - r_{f,t}) + \beta_2 SMB_t + \beta_3 HML_t + \epsilon_{i,t}$$

where $r_{i,t}$ is the total return of a stock or portfolio *i* at time *t*, $r_{f,t}$ is the risk free rate of return at time *t*, $r_{m,t}$ is the total market portfolio return at time *t*, $r_{i,t} - r_{f,t}$ refers to the expected excess return, $r_{m,t} - r_{f,t}$ is the excess return on the market portfolio, SMB_t is small minus big (size premium), HML_t is high minus low (value premium), $\beta_{1,2,3}$ is factor coefficients and $\epsilon_{i,t}$ is residuals of the regression model.

The factors used, in addition to the market factor, are SMB and HML. The SMB factor includes publicly traded companies with small market capitalization that generates higher returns. The factor

shows the difference between the returns on small cap and large cap portfolios with the same weighted average book-to-market ratio. HML includes value stock with a high book-to-market ratio that generates excess return compared to the market. The factor is the difference between returns on high and low book-to-market equity portfolios with about the same weighted average size (Fama and French, 1992). SMB and HML are estimated using the following equations (French, 2020):

- **SMB (Small Minus Big)** – the average return on three small portfolios minus the average return on three big portfolios:

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

- **HML (High Minus Low)** – the average return on two value portfolios minus the average return in two growth portfolios:

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

Although the three-factor model have been a widely used tool in research, the model has met some criticism. Fama and French (1993) look at average returns on certain factors as evidence of expected returns. Black (1995) states that this will normally give a highly inaccurate estimate as it is assumed that the expected return in excess of the riskless rate is held constant. According to Schwert (2002), the estimates of factor risk premiums seems very high, particularly the book-to-market factor. If the book-to-market premium is too high, the returns vary too much with the risk factor. From this perspective, the evidence that the three-factor model is a good linear model of risk and return could be an accidental description of an anomaly (Schwert, 2002; MacKinlay, 1995). The research conducted on the SMB and the HML shows deviating results. The SMB factor was the strongest in the beginning of the 1980s but hasn't been significant since the end of the beforementioned century. However, the HML factor is proved to be significant. The factor is fundamentally risky, and performs bad in times of recession, i.e. a strategy of investing in portfolios consisting of value stocks tends to perform poorly (Ang, 2014).

3.3.1 Independent Variables

- **Alpha, α_i,** is used to capture the variation in stock returns that cannot be explained by the sensitivity of the three coefficients to fluctuations in their respective risk factors.
- **Market beta, β_1,** is a proxy for the sensitivity of stock returns to market fluctuations and corresponds to the variable used in the CAPM. The only major difference in Fama and French (1992) model is that this factor is not the only explanatory coefficient for stock returns.
- **Coefficient β_2,** size loading factor.
- **Coefficient β_3,** value loading factor.
- **The SMB factor**, the size premium, reflects the risk factor that is related to the market value of the companies. According to Fama and French (1996), small businesses may experience longer periods of "revenue depression" that doesn't affect large companies in the same extent. Fama and French (1992) and Banz (1981) found that this factor has a significant degree of explanatory power regarding variances in stock returns across different sized companies, which could imply that companies pays a size premium.
- **The HML factor,** the value premium, shows that companies with a low stock price relative to book value generates lower income on their assets, while companies with a high stock price relative to book value generates higher income (Fama and French, 1992). The BE/ME ratio represents a proxy for a risk factor that should be taken into account when explaining variations in stock returns.
- **Rm-rf** is a proxy for the market factor in stock returns and represents the market risk premium. Rm is the return on the market portfolio and rf represent the risk-free rate.

3.4 Carhart Four-factor Model

Carhart (1997) built on Fama and French's (1993) three-factor model by introducing an additional factor. The factor is intended to capture Jegadeesh and Titman's (1993) one-year momentum anomaly. The momentum factor shows the tendency for a given asset to continue on a certain path, such as falling or rising. The additional factor is based on designing portfolios which consists of buying stocks that have performed well over time and selling stocks that have performed poorly. The factor is then added to the Fama & French three-factor model and is expressed as *winners minus losers* (WML). Carhart four-factor model is illustrated in the equation below:

Equation 5: Carhart four-factor

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_1(r_{m,t} - r_{f,t}) + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 WML + \epsilon_{i,t}$$

The model may be interpreted as a performance attribution model. The coefficients and the premia on the factor-mimicking portfolios indicates the proportion of mean return that is attributable to four elementary strategies. These four strategies consist of high versus low beta stocks, large versus small market capitalization stocks, value versus growth stocks, and lastly one-year return momentum versus contrarian stocks (Carhart, 1993). The equation for WML are estimated using the equation below:

- **WML (Winners Minus Losers)** – the equal-weight average of the returns for the two winner portfolios for a region minus the average of the returns for the two loser portfolios:

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

3.4.1 Independent Variables

Carhart four-factor model includes the same independent variables as Fama & French three factor model. In addition, one extra factor is included:

- **Coefficient $\boldsymbol{\beta_4}$,** loading factor for WML
- **WML,** is a cross-sectional momentum factor that improves the explanatory power of the multifactor model and is the premium on winners minus losers. Carhart (1997) states that covariance with the WML factor appears to be a good indication of the momentum of the underlying stocks in a portfolio.

3.5 Fama & French Five-factor Model

As studies found the Fama and French (1993) three-factor model to be inadequate, Fama and French (2015) added two new factors to the model. The five-factor model is focused on capturing the size, value, profitability and investment patterns in average stock returns. They express that the model explains between 71% and 94% of the cross-section variance of expected returns for the *size*, *Book-to-market ratio (B/M)*, *profitability (OP)* and *investment* (*INV)* portfolios. The equation for Fama & French five-factor model is illustrated below:

Equation 6: Fama & French five-factor

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_1(r_{m,t} - r_{f,t}) + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 RMW_t + \beta_5 CMA_t \epsilon_{i,t}$$

The profitability factor, *robust minus weak* (RMW), is constructed by calculating the difference between returns on diversified portfolios of stocks with strong operating profitability and weak operating profitability. The investment factor, *conservative minus aggressive* (CMA), is constructed by calculating the difference between returns on diversified portfolios of stocks of low and high investment firms. RMW and CMA are estimated using the following equations (French, 2020):

- **RMW (Robust Minus Weak)** – the average return on the two robust operating profitability portfolios minus the average return on the two weak operating profitability portfolios

$$RMW = \frac{1}{2}(Small\ Robust + Big\ Robust) - \frac{1}{2}(Small\ Weak + Big\ Weak)$$

- **CMA (Conservative Minus Aggressive)** – the average return on the two conservative investment portfolios minus the average return on the two aggressive investment portfolios

$$CMA = \frac{1}{2}(Small\ Conservative + Big\ Conservative) - \frac{1}{2}(Small\ Aggressive + Big\ Aggressive)$$

The model suggests that HML is a redundant factor in the sense that its high average return is fully captured by its exposures to $r_{m,t} - r_{f,t}$, SMB, RMW and CMA. In the scenario where the sole interest is abnormal returns (measured by regression intercepts), a four-factor model that omits HML performs as well as the five-factor model. However, if the analysis focuses on portfolios which tilt toward size, value, profitability and investment premiums, the five-factor model may be the more appropriate choice (Fama and French, 2015)

A problem with the five-factor model is its inability to capture low average returns on small stocks, where the returns behave in a similar pattern to those of firms that invest heavily despite low profitability (Fama and French, 2015). The new added features of the model have received some criticism, as the robustness and intuition the model represents have been questioned. In addition, it ignores the, by now widely accepted, momentum effect (Blitz et al., 2018).

3.5.1 Independent Variables

Fama & French five-factor model builds upon the three-factor model by adding two additional factors:

- **Coefficient β_4,** loading factor for RMW.
- **Coefficient β_5,** loading factor for CMA.
- **RMW,** is the profitability factor, and shows the difference between the returns of firms with robust (high) and weak (low) operating profitability.
- **CMA,** is the investment factor, and is the difference between the returns of firms that invest conservatively and firms that invest aggressively.

3.6 Portfolio Diversification

Investors are being exposed to different type of risks when buying stocks and other securities. The finance literature distinguishes between systematic and unsystematic risk, where systematic risk covers a grouping of factors that are common to all shares. This refers to the risks that could be caused by financial system instability, macroeconomic development, idiosyncratic events and other interdependencies in the overall market. Systematic risk, or non-diversifiable risk, is inherent in nature and not controlled by an individual or group, hence makes it nearly impossible to mitigate the risk involved. Unsystematic risk, on the other hand, are associated with a specific industry or segment and is manageable by means of an intelligent portfolio composition, which means that the risk can be reduced by diversification (Bodie, Kane and Marcus, 2017).

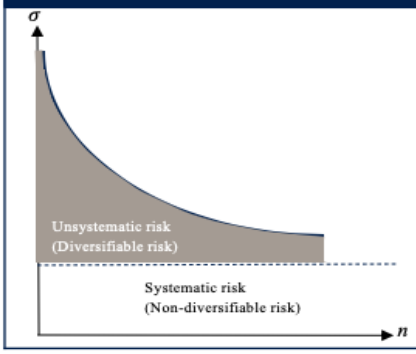
	Ex. of systematic risk	Ex. of unsystematic risk
	Interest rate risk	Business risk
	Inflation risk	Financial risk
	Market risk	Technological development
	Exchange rate risk	Lawsuits
	War	Management changes
	Epidemic	Capital investment decisions
	Government legislation changes	Entry of competitor

Figure 4: Portfolio risk (source: Bodie, Kane and Marcus, 2017). Authors own drawings.

By investing in a specific stock, you are exposed to the company's future prospects and performance. The return on investment is therefore dependent on the underlying risk for the company. In a portfolio context, it is possible to diversify large parts of the unsystematic risk so that the investor becomes less dependent on the development of individual companies. The relationship between the portfolio risk (σ) and the number of shares (n) is illustrated in figure 4.

As the unsystematic risk is diversifiable, investors should not expect to be rewarded for taking on this type of risk in form of a risk premium. The market prices the risk premium based on the systematic risk, where investors are compensated accordingly. The greater an asset's covariance with the other assets in the portfolio, the more it contributes to portfolio variance. The portfolio diversification is of value as long as assets are less than perfectly correlated, i.e. an asset that is perfectly negatively correlated with a portfolio can serve as a perfect hedge. The perfect hedge asset can reduce the portfolio variance to zero (Bodie, Kane and Marcus, 2017).

3.7 Key Takeaways

Factor models in the analysis:

- The relationship between ESG and return will be analyzed by applying multiple factor models. Gradually including additional factors, may help develop an understanding of their separate effect. They will capture the return that is due to the portfolio's exposure to the respective factors.
- Alpha will then represent the return which may not be explained by the exposure to systematic risk factors.
- The following factor models will be used in the empirical analysis:
  - CAPM
  - Fama & French three-factor
  - Carhart four-factor
  - Fama & French five-factor

4 Dataset

This chapter will present the dataset that have been studied in the analysis. This includes the market portfolio, risk-free rate, financial figures, in addition to the ESG and controversy score. A thorough explanation of the ESG and controversy score is presented, where both the ESG data provider and how the different scores is calculated is illustrated. Next, the process of the self-constructed factors and how these are calculated is described. Finally, the data preparation is accounted for.

For this study, publicly available data will be used in order to answer the research question. With respect to the focus on corporate responses to e.g. environmental reporting, labor rights and management compensation, this study only considers activities that have been published and thus reported in corporate reports or news sources. According to Ziegler et al. (2011), this approach is adequate since information about the real extent of activities beyond publicly available information is not available for the investors on the stock market. This analysis is therefore particularly based on general theoretical arguments between disclosed corporate ESG activities and economic success. The paper assumes that the market is efficient in semi-strong form, which means that companies that are included in the analysis are selected on the basis that they are publicly traded and that the market incorporates all publicly available information in the share price. This includes that the effect of ESG specific information already have been accounted for in the share price.

4.1 Market Portfolio

It is not possible to collect an international market portfolio that covers the entire European market. For this reason, it has been necessary to use a proxy for the market index. In order to estimate the market risk premium, a broad market index is used, which is assumed to be a proxy for the entire market. The chosen proxy is the Thomson Reuters Europe 500 Index, which have been used as market proxy for the study perfromed by Ziegler, Busch and Hoffman (2011). Another representative market index that could have been used is the Stoxx Europe 600. This index has been used in the study conducted by Oberndorfer et al. (2013) but have not been used for this analysis as the data only goes back to April 2004 and thus doesn't cover the entire relevant period. Thomson Reuters Europe 500 index is a market capitalization weighted standalone index, derived from Thomson Reuters Europe Index, and is collected through Thomson Reuters Datastream. The index includes Euro Zone

countries and a few other European countries. The 500 largest stocks by market capitalization are then selected to create a Europe 500 index and is rebalanced with the Thomson Reuters Equity Indices (Thomson Reuters, 2015). The total return index - in the same way as the market values – have been consistently denominated in Euros and on a monthly basis. This is a procedure that is common in financial studies (e.g. Ziegler et al., 2011; Oberndorfer et al., 2013) due to the international investor perspective. The total return index corrects for dividend payments, and the correlation is sufficient to ensure that the results are not adversely affected by dividends.

4.2 Risk-free Rate

Government bonds are a usual proxy for risk-free rate. When studying companies form multiple countries, a common risk-free rate is challenging to decide on in the myriad of government bonds. As this study concerns European companies where all the prices collected are denominated in Euro, the risk-free rate that is used is based on 1-month Euro Interbank Offered Rate (EURIBOR). This is the average interbank interest rate at which European banks are prepared to lend to one another within the Euro zone. The EURIBOR is updated on a daily basis and is considered as the benchmark for the euro zone money market (Euribor, 2020). In the literature, Ziegler et al. (2011) and Oberndorfer et al. (2013) uses the 1-month Euro Interbank Offered Rate (EURIBOR) as a proxy for their risk-free rate. This rate is therefore evaluated to be the most relevant for this study.

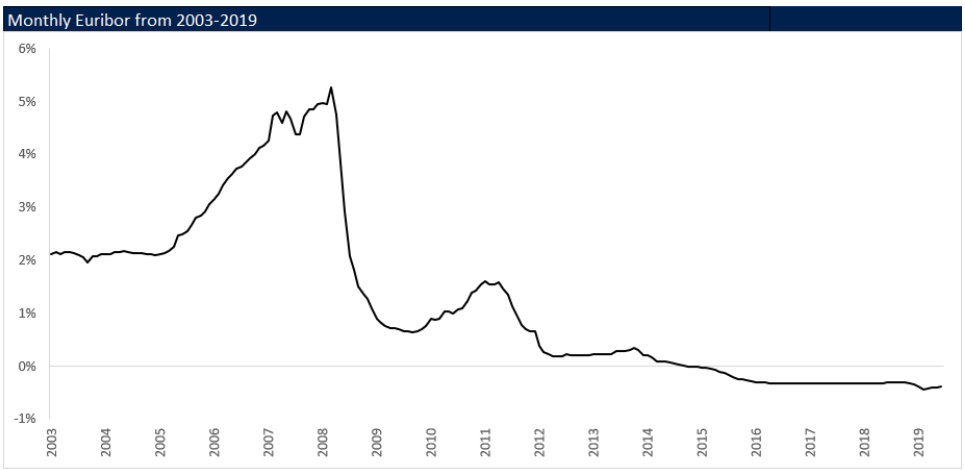


Figure 5: 1- Month EURIBOR from 2003-2019. Source: Thomson Reuters.

Figure 5 illustrates the development of the 1-month EURIBOR. It clearly shows how the rate rises to the beginning of the financial crisis in 2008, and then fall to current historically low levels.

4.3 Financial Figures

The monthly return is based on the change in the Return Index (ΔRI_t) collected from Thomson Reuters Datastream, and includes dividend reinvestment. The return index is calculated using the following equation:

Equation 7: Return Index

$$RI_t = RI_{t-1} * \frac{P_t}{P_{t-1}}$$

The only exception is on the reinvestment date of the dividend, where RI_t is expressed as:

Equation 8: Return Index with dividend

$$RI_t = RI_{t-1} * \frac{P_t + D_t}{P_{t-1}}$$

P_t is associated to the price on day *t*, and D_t is the dividend on the reinvestment date *t*. Finally, the return index is converted to return expressed as percentage points (r_t):

Equation 9: Return Index (percentage points)

$$r_t = \Delta RI_t = \frac{RI_t - RI_{t-1}}{RI_{t-1}} * 100$$

The monthly return reported in the end of each month is used during the data processing.

The risk-factors SMB, HML, WML, RMW and CMA is collected through "Fama/French European 5 factors" and "European Momentum Factor" from Kenneth French database, which is gathered from Bloomberg (French, 2020). The data is reported as monthly figures, which are updated versions of the figures used in Fama and French (2012). Retrieving the factors from the Kenneth French database improves the comparability and reliability of the results, as this is common practice for factor model analyses (Friede, Busch and Bassen, 2015).

The monthly market value (MV) for each company has been gathered in order to calculate the weighted return for value-weighted portfolios with different ESG-scores, which corresponds to the same approach as Halbritter and Dorfleitner (2015). These numbers have been collected through Thomson Reuters Datastream. MV from the end of June year *t* is used to weight companies under the composition of portfolios for the period July year *t* until June year *t+1*. MV from the start of the

month is used in order to weigh the portfolio's monthly return. As the monthly returns needs to be comparable between firms, all numbers are converted into a single currency. Euro is the most used currency in Europe and have thus been evaluated to be the most representative currency, aligned with studies made by Ziegler et al., (2011) and Oberndorfer et al., (2013). This have been conducted using a built-in function in Thomson Reuters Datastream.

4.4 Measures for Responsibility

Several alternative methods in the literature review were identified for measuring the relationship between responsible investments and financial performance. In recent studies, ESG ratings from different rating agencies are increasingly being used. Escrig- Olmedo et al. (2019), made an analysis of the most representative ESG rating and information providers in two time periods, 2008 and 2018 respectively. They conclude that ESG rating agencies are contributing to sustainable development with the increased focus on sustainability principles.

The practice of ESG reporting is not standardized, and companies reports in various formats, units and scales. Some companies disclose responses to climate change and other relevant ESG data as a subpart of their quarterly and annual reports, often based on the UN Global Compact 17 sustainable development goals. This may include sustainable investments and initiatives, with targeted numbers for responsibility for the upcoming years. However, some companies do not provide investors with relevant ESG data. The rating institutions have thus been given a decisive role in disseminating ESG data, and the ratings are increasingly being used as universal proxies for corporate responsibility and social performance. As shown in previous studies, ratings can be used as a quantitative measure of corporate ESG commitment for companies, and thus makes it relevant to use for this study.

4.4.1 Data Providers

Several different providers of ESG rating exists. A common feature for every data provider is that they analyze company performance on the parameters that include environment, social and governance issues, and reveals quantified or qualitative scores for each of them. Their goal is to integrate sustainability and financial aspects by identifying non-traditional sources of information in order to contribute to the overall risk evaluation for investors. Throughout the research process, the authors have observed several different providers of ESG data that have been widely used in the literature.

Kinder, Lydenberg, Domini & Company (KLD) was founded in 1989 and offers institutional investors easy-to-use social responsibility research on companies. The KLD have been acquired by Morgan Stanley Capital International and are better known as the MSCI KLD 400 Social Index. KLD does not offer numerical values for ESG score but use binary indicators for strengths and concerns. The platform has compiled socially responsible indices where they manage the screening process, which makes it possible for the investor to invest in already screened companies (MSCI, 2020).

Bloomberg ESG data collects ESG data from more than 13.000 companies in 83 countries. Their data provides up to 12 years of annual ESG data and daily granular governance data back to 2013 for nearly 4.500 companies. The data are accessible through the Bloomberg Terminal, and displays it alongside fundamental financial data (Bloomberg, 2020).

Innovest was founded in 1995, but was in 2010 acquired by MSCI, and thus integrated in the KLD data. For this reason, Innovest have not been mentioned in the literature since 2010, and their data is only available through KLD (MSCI, 2020).

Asset 4 was founded in 2003 and was acquired by Thomson Reuters Inc. in 2009. Asset 4 provides objective, comparable and auditable extra-financial information which is quantified and made accessible through Thomson Reuters Datastream. The data provides an overall score for companies, as well as individual scores on the four pillars that it is based on. These includes an overall weighted rating score, corporate governance score, social score and environmental score (Thomson Reuters, 2020).

Refinitiv was founded in 2018, and is jointly owned by BlackRock Group LP (55%) and Thomson Reuters (45%), and is a global provider of financial market data and infrastructure. Thomson Reuters Datastream does now include the solutions provided by Refinitiv as key enhancements over the legacy equal-weighted Asset 4. These includes, among others:

- ESG controversies overlay in order to magnify potential impact of controversies on the overall ESG rating score.
- Industry and country benchmarks to facilitate comparable analysis within peer groups
- Percentile rank scoring methodology to eliminate hidden layers of calculations.

Refinitiv offers one of the most comprehensive ESG databases in the industry, with history going back to 2002. It covers over 70% of global market capitalization across more than 400 different ESG metrics (Refinitv, 2019).

4.4.2 Motivation for used ESG Proxy

In the literature review, there was identified several studies that applied data from KLD (Sharfman and Fernando 2008), (Galema et al., 2008), (Rennebog et al., 2008) (El Ghoul et al. 2011), (Attig et al., 2013), and Asset4 (Ziegler et al., 2011), (Ioannou and Serafeim 2012), (Bouslah et al. 2013), (Halbritter and Dorfleitner 2015), (Aouadi and Marsat, 2018), (Chollet and Sandwidi 2018), (Blackrock, 2018). Results from previous research show that there exists a big difference between the methodology and rating criteria from the individual providers of ESG data. Dorfleitner et al. (2014) explains that there is clear difference in the methodology, which in turn yields significant differences between the scores. In addition, the different providers are only to a small degree transparent in the way they reveal their individual ratings and there is reason to believe that there exists some kind of bias in the ESG scoring. Their study rejects the hypothesis that KLD, Asset 4 and Bloomberg investigates the same parameters. Halbritter et al. (2015) find the similar results throughout their research. KLD was one of the first ESG rating providers and goes back to early 1990s and is viewed as transparent in regard to strengths and concerns for the individual company. For this reason, KLD data is used in numerous empirical papers. However, KLD does not provide the user with numerical values, where the data have to be transformed to numerical values in order to be compatible with the scores of Asset4 and Bloomberg. This can be done by following Kempf and Osthoff (2007) method by reverting all concerns back into strengths by using the opposite binary value. The results from Halbritter et al. (2015) shows that the scores diverge on several parameters as unequal weightings are being used, which emphasizes that the method of quantification of ESG data depends on a subjective basis.

The choice of rating proxy can lead to dissimilar results when analyzing the relationship between ESG and financial performance. Although many studies have been conducted by using KLD data, many investors turn to alternative proxies, such as Asset4, for collection of ESG data. In 2010, significant changes were made to KLD's methodology (Dorfleitner et al., 2014), which makes it difficult to compare results from the period before and after 2010. The same problem does not occur with the use of Asset4. The literature that identifies the relationship between the scores and financial

performance is limited. Halbritter et al. (2015) and Dorfleitner et al. (2014) emphasizes that there is a significant uncovered area in the literature in using alternative scores as proxies, as KLD scores are not assumed to be directly transferable to scores from Asset4 and other rating institutions that provide quantitative scores.

As discussed, there exists some mayor differences between the most used data providers. The analysis of this thesis will use ESG scores and controversy scores provided from Refinitiv, which is the successor of the mentioned Asset4. The main reasons for this choice are that Refinitiv offers data in the entire investigated period of research, it includes both ESG and controversy scores, all data are numerically assessed and most importantly, it is widely used from previous empirical studies. The next section will elaborate on Refinitiv and its enhancements relative to its legacy equal-weighted Asset4. Furthermore, how the ESG scores and controversy scores are assessed will be discussed. Finally, the self-constructed factors for ESG and controversy scores will be described and how the data preparation were done.

4.5 Refinitiv ESG Scores

With over 150 content research analysts, Refinitiv processes vast volumes of public information with a purpose of providing relevant and objective ESG data. It provides the user with a worldwide coverage of companies listed on the S&P500, MSCI World Index, MSCI Emerging Markets, DAX, FTSE250 and STOXX600. The data is benchmarked against Thomson Reuters Business Classifications and is made available through Thomson Reuters Datastream. Refinitiv offers one of the most comprehensive ESG databases in the industry and is a successor to the existing Asset4 (Refinitiv, 2019). Critical differences are expressed in table 2.

	Refinitiv	Asset 4
Scoring methodology	Percentile rank	Z-scoring
Industry benchmarks	Yes	No
Country benchmarks	Yes	No
Industry-specific weights	Yes	No
Industry-relevant measures	Yes	Yes
Standard pillar scores (E,S,G)	Yes	Yes
Inclusion of controversy score	Yes	No

Table 2: Differences between Refinitiv and Asset 4.

Based on the arguments presented from previous section, the immediate benefit of using Refinitiv as a proxy for environmental performance, is that it measures data input from over 400 different ESG indicators across 4 main categories (Refinitiv, 2019). In comparison, KLD uses only 80 parameters as their data input (Dhaliwal, Li, Tsand, and Yang, 2014). The ratings provided by Refinitiv is therefore a fairly detailed proxy for the underlying ESG performance for companies. There are extensive studies about KLD and Asset 4, while the use of the new Refinitiv database are poorly represented. In addition, this study will be supplemented with the newly introduced Controversy Score. The literature review indicates that there hasn't been written much about the use of controversy as a separate score or assessment criterion on the same basis as the ESG-score. It is further expected that the results from this study will contribute to a more nuanced coverage of the relationship between ESG and financial performance by using a fairly new proxy for the data collection.

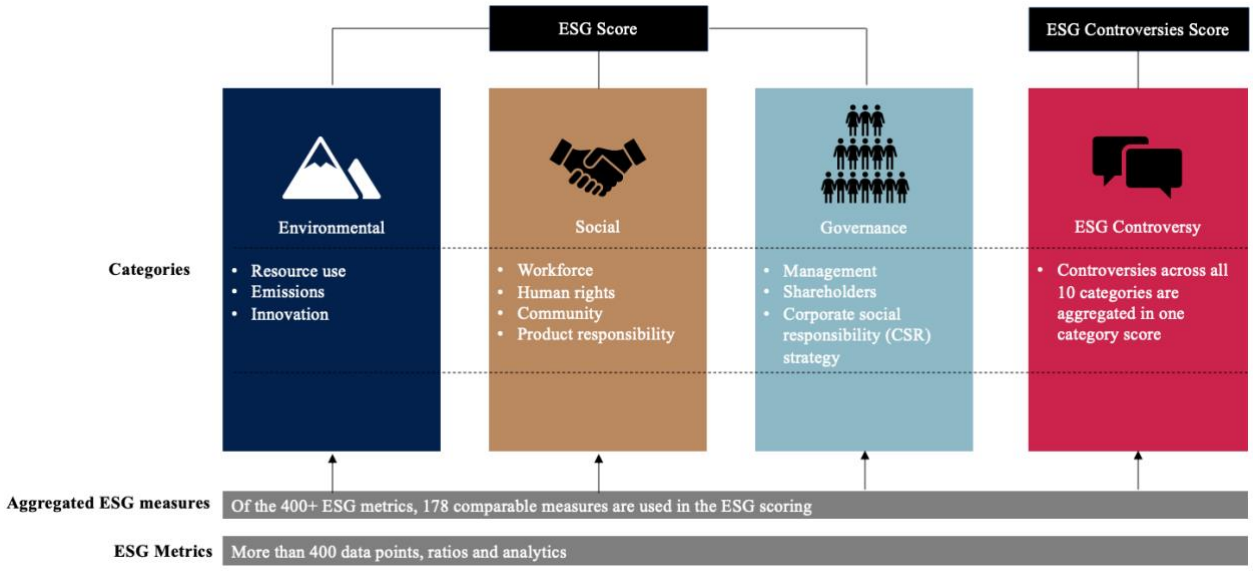


Table 3: Overview of Refinitiv ESG model. Authors own drawings.

4.5.1 ESG Score

The ESG universe in the database from Refinitiv consists of more than 7,000 companies globally and 1,200+ in Europe. The ESG Scores are designed to transparently and objectively measure a company's ESG performance across 10 main themes (Refinitiv, 2019). These categories are based on company-reported information and can be viewed in table 4.

	Category	Score
Environment	Resource use	19
	Emissions	22
	Innovation	20
Social	Workforce	29
	Human rights	8
	Community	14
	Product responsibility	12
Governance	Management	34
	Shareholders	12
	CSR strategy	8

Table 4: Overview of categories evaluated in the ESG-scores

The scores presented on the right side in table 4, can be both qualitative and quantitative. Qualitative metrics are Boolean questions and the values are "Yes" or "No" and then converted to a numerical value. "Yes" is assigned the value of 1 and "No" the value of 0,5. Some indicators are industry-specific and not relevant for all companies. The indicator is then excluded from the calculation and will not be relevant, expressed as "N/R". If a company lacks data for a quantitative metric, the potential score is removed from the calculation. The only exception is greenhouse gas emissions, where an estimated emission is accounted for when it's not reported by the companies. If a company receive a positive or high value on a negative attribute, the score will be treated as a negative number (Refinitiv, 2019).

A percentile score is assigned to each company by using the formula in equation 10. The score is calculated for each relevant ESG metric for the relevant company, i.e. "estimated CO2 equivalents emission"," policy emissions" and "targets emissions". The score that is produced for each metric is based on ranking, where each company's rating depends on how many companies that have lower value, equal value and/or if the companies report a value at all. Benchmarks for environmental and social conditions, and also the controversy score, are based on relevant industries, where data is collected from Thomson Reuters Business Classification. For the governance categories, country of headquarters is the applied benchmark (Refinitiv, 2019).

Equation 10: ESG percentile score

$$Score = \frac{\#\ of\ companies\ with\ a\ worse\ value + \frac{\#\ of\ companies\ with\ the\ same\ value\ as\ company\ x}{2}}{\#\ of\ companies\ with\ a\ value}$$

Based on the nature of the metric, relevant numeric values are assigned and calculated, where the percentile score calculation formula is applied for each measure. This is illustrated in equation 10.

Identical procedures are applied every data point in the category. When the percentile scores are calculated data-point wise, average scores are developed for individual companies and sorted from highest to lowest. The different category weights are determined by the number of indicators that make up each category. Resultingly, categories like management, e.g. diversity, committees and compensation, which comprise parameters with higher levels of transparency will be weighted more heavily than less reported categories, such as CSR strategy. After the calculations are completed, each company will e.g. receive an average emission category percentile score and an emission category score. Lastly, a grade for each company (ranging from A+ to D-) is presented (Refinitiv, 2019).

4.5.2 ESG Controversy Score

The ESG Controversy Score is calculated based on 23 ESG controversy topics. Controversies denote scandals that involves the respective company. The impact of the relevant event can also be seen in the following year if there are new developments related to the negative event. Examples of such events could be ongoing legislation, lawsuits, fines and disputes. All new articles that shows up in the aftermath of the event are captured and are accounted for in the score (Refinitiv, 2019).

An ESG Controversy score is valid for one fiscal year. If a controversy occurs before a fiscal year is ended, the event is included in the latest completed fiscal year's score. This secures consistently updated scores for each company. For instance, the latest completed fiscal year is December 31st, 2019. There exists one controversy on May 15th, 2020 and one on Sep 15th, 2020, i.e. before the fiscal year 2020 is completed. Both of the controversies are accounted for under recent controversies and are thus included in the scoring for fiscal year 2019. When fiscal year 2020 is completed, the mentioned controversies will be accounted for in fiscal year 2020. This means that the data for last year may change continuously as controversies occur and are also dependent on when the controversy score for next year fiscal year is published (Refinitiv, 2019).

4.5.3 Self-Constructed Factors for ESG and Controversies

In order to construct the ESG factor and the Controversy factor, the companies are first sorted after their respective scores ranging from highest to lowest, for each year. Furthermore, the companies are divided into three portfolios based on the value of the companies, with splits on the 30th and 70th percentile. The factors are based on lagged scores, which means that the portfolios used to make the factors from July year *t* until June year *t+1* are based on scores from fiscal year ended in year *t-1*. This is to account for associated time effects of scoring. These portfolios are rebalanced in the end of June for each consecutive year. The value-weighted return for both the high and low portfolios are then calculated, where all stock returns are weighted according to their market capitalization. Both the ESG and Controversy factor are then found by subtracting the value-weighted return of the low portfolio from the high portfolio.

4.6 Data Preparation

The dataset is collected from Refinitiv through Thomson Reuters Datastream and consists of approximately 7,000 companies from the main stock exchanges worldwide. In the very beginning of the collection of the ESG data, 1,745 European companies were registered. The companies that did not have necessary data, were removed from that respective period. Next, after the missing observations were removed, the return and MV was *winsorized* according to Fama and French (1992). Winsorizing is the transformation of statistics by limiting extreme values in the statistical data, with the purpose of reducing the effect of possible spurious outliers. By removing observations at the 0.5% and 99.5% level, the mean, standard deviation, skewness and kurtosis are reduced to a level that compares better to return series over a long period of time (Lipson, Mortal and Schill, 2011). Finally, after the corrections have been made, the dataset consists of 1,254 companies, and corresponds to what Thomson Reuters lists as their coverage of European companies in their database. The companies represent different sectors and industries, and thus makes the sample representative. Data is subtracted from 31.07.2003 to 31.06.2019. The data used for the analysis consists of revised and published figures from the individual companies. This means that there have been several independent parties approving the reliability of the figures. It is assumed that the liquidity of each individual share is sufficient in order for the total return to be representative during the period of the analysis. Schmidt et al. (2011) shows that data from Thomson Reuters is well suited to construct high-quality, replicable portfolios and risk factors to be tested for in Carhart four factor and Fama & French

three- and five factors models, so that reliable calculations can be made for hypothesis and research questions.

4.7 Key Takeaways

Subject	Data
Market portfolio	Thomson Reuters Europe 500 Index
Risk-free rate	1-month EURIBOR
Monthly variables	Market Value (MV) and Return Index (RI)
Database	Thomson Reuters Datastream
ESG score provider	Refinitiv
Controversy score provider	Refinitiv

5 Methodology

The first part in this chapter will elaborate on the motivation behind the choice of methods used for the analysis. The second part will focus on which methods that have been used to analyze the portfolios, and how these are constructed. The final part will present the model testing, where several tests of assumptions for multiple regressions will be described.

5.1 Motivation for used Methods

Existing financial literature proposes two main approaches to investigate the relationship between ESG and financial performance. One approach is to aggregate stocks into portfolios and use them as base assets. The portfolios are constructed based on specified parameters such as share of total population, market value or high/low performers on specific criteria. When the portfolios are designed as time series, it is possible to apply factor models. The second approach is to structure individual stocks as panels. By studying panel-based company data, one can investigate the direct effect of ESG on a company's return. Portfolio studies may miss details in the company-specific information.

Constructing ESG portfolios is one of the most common approaches for investigating the relationship between corporate responsibility and financial performance (Halbritter and Dorfleitner, 2015). Ang, Liu and Schwarz (2017) find that portfolios provide less uncertain factor exposure, which should translate into more precise estimates and possibly lower standard errors for risk premia. This is because by grouping companies in portfolios, a substantial amount of the information on the individual factor exposure is diversified. Blume (1970) stated this as the original motivation for performing portfolio analyses. Later, numerous authors such as Black, Jensen and Scholes (1972), Fama and MacBeth (1973) and Fama and French (1993), used the same motivation to choose portfolios as base assets.

However, as portfolio analyses also risk efficiency losses compared to an analysis focusing on individual companies, it would be insightful to apply both. Moreover, in addition to the benefits with the respective methods, the analysis' robustness would also increase by exploring more than one method. However, due to limited data access at CBS facilities, this was not possible.

5.2 Portfolios

5.2.1 Portfolio Construction

The dataset, after initial cleaning, consists of 1,254 companies with their own unique ESG score and Controversy score. The scores are collected on a yearly basis, from the end of July 2003 until the end of June 2019. For year *t*, ESG scores from the financial year ended in year *t-1* are used to take into account a delayed effect. This means that the ESG score that is used is available at the time the portfolio is compiled and can be used in practice.

There is no consensus in the literature concerning the most appropriate cut-off levels for the portfolios. For example, Kempf and Osthoff (2007) uses 10% as their base strategy and continues with performing additional checks with larger subsets of the selected sample. The high-ranked (low-ranked) portfolios made by Derwall et al. (2005) consists of companies making up the 30% of total selection capitalization. Furthermore, Derwall et al. (2005) chooses to divide the portfolios into 20% of the highest (lowest) and 40% of the highest (lowest) for their robustness test. Halbritter and Dorfleitner (2015) makes portfolios that consists of only 1% of the selection, and then shifts the threshold to create larger portfolios. The approach for this study will be to use three different cut-off levels. Portfolios are divided into 20% of the highest (lowest) performers for both ESG scores and Controversy scores, which will be considered as the main strategy. Next, cut of-levels of 10% and 50% will be used in the robustness tests.

The first portfolio composition is performed in the end of July 2003 and is then kept for one year before being rebalanced at the end of June 2004. This procedure is performed for each consecutive year until the end of June 2019. The monthly value-weighted return is measured at the end of each month. By comparing the return on these portfolios against one another it is possible to compare returns from portfolios with different scores. Management costs and certain composition restrictions that real portfolios may be exposed to are disregarded in the analysis.

5.2.2 Pricing Models

In order to control the various portfolios against known risk-premiums, timeseries regressions were performed by gradually including more factors. When the factors are gradually added, it is easier to obtain a better understanding of the effect of the various factors and how they affect each other. The idea is that the factors will capture the return that is due to the portfolio's exposure against the relevant

factors. Alpha will then illustrate the return that cannot be explained by the exposure to the systematic risk factors. By controlling for the systematic risk factors in this way, i.e. comparing alphas with different ESG and Controversy scores, it will to a greater extent be possible to isolate the actual effect of ESG and Controversy towards returns.

The first pricing model is the CAPM, as shown in chapter 3.2. Alpha then gauges the performance of an investment against a proxy, in this case the entire selection, that is considered to represent the market's movement as a whole, where the alpha consists of the return that cannot be explained by the exposure to the market. Subsequently, the portfolio's size and value effects are investigated. This is done by including SMB (small minus big) and HML (high minus low), as shown in section 3.3. If the constructed portfolios yield different returns due to exposure to either SMB or HML, the effect will be that alpha moves towards zero by including the factors in the regression. The next pricing model includes WML (winners minus losers), as shown in section 3.4. This factor considers the momentum effect, which is the tendency for assets to continue on a given path between two periods, e.g. rising or falling. CMA (conservative minus aggressive) and RMW (robust minus weak), as shown in chapter 3.5, are added to the regression equation to fulfil the Fama & French 5-factor model. When these factors are included, it is easier to assess whether any of the return differences are due to the ESG or Controversies.

Lastly, the self-constructed factors that represents ESG and Controversies are added, as illustrated in equation 11. These last two factors are added to control for the covariance between the different scores, in order to check if these factors could yield some consistent notable differences in the results of the regressions. This could be e.g. that if a portfolio with a high controversy score has a significant alpha after the ESG factor is included, the controversy score may generate excess return that cannot be explained by differences in ESG scores.

The relevant regressions are variations of equation (11). If a model contains factors for which the portfolios do not have significant exposure, it may be omitted in subsequent models in order to not lose degrees of freedom in the model.

Equation 11: Fama & French five-factor and self-constructed factors

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_1(r_{m,t} - r_{f,t}) + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 RMW_t + \beta_5 CMA_t + \beta_6 ESG_t + \beta_7 Controversy_t + \epsilon_{i,t}$$

5.3 Model Testing

The various models applied in the analysis are based on a set of assumptions. To justify the results from the analysis, it is necessary that the data satisfies these. Firstly, linearity was tested for all datasets in the regressions by plotting the residuals against the independent variables. When the linearity assumption is satisfied, the residual plots will be evenly distributed along a horizontal line without any systematic variations. After visual inspection of the plots, the authors conclude that the data satisfy the assumption of linearity (Appendix, 12.5).

To assess potential heteroskedasticity in the time series regressions, Breusch- Pagan (1979)/ Cook-Weisberg (1983) test was applied, which assumes normally distributed residuals (appendix, 12.6). More thoroughly, the test assesses whether the variance of the residuals is dependent on the independent variables. Heteroskedasticity is present when the standard errors of a variable are non-constant when measured over a specific time interval. After assessing the test output for the regressions that did not satisfy the assumption of normally distributed residuals, residual plots against estimated values were further examined visually for traces of heteroskedasticity. If the residual errors tend to fan out over time, it may be proof of heteroskedasticity. For the regressions where heteroskedastic problems were identified, White standard errors, which are robust to this effect, were calculated and applied. However, the robust standard errors did not alter the regression output substantially.

Autocorrelation was investigated through the Breusch-Godfrey (1978) test for higher-order autocorrelation (appendix, 12.6). Unlike the Durbin-Watson (1992) test, Breusch-Godfrey (1978) do not assume that all variables are strictly exogenous, nor is it limited to testing first-order autocorrelation only. If a time series regression had problems with autocorrelation, Cochrane-Orcutt (1949) estimates were applied to counter this effect.

To avoid spurious relationships, the data must be stationary. Since returns are at the center of the study, it could be reasonable to expect some degree of stationarity in the time series. However, this has formally been confirmed by applying a Dickey-Fuller (1979) test, alternatively an augmented Dickey-Fullers test, where it is possible to further identify problems with autocorrelation. A stationary time series is one whose statistical properties are constant over time. This is not to be confused with no change, but that the *way* it changes does not itself change over time. As seen in Appendix 12.6 the authors have not identified challenges with stationarity in either of the models.

Further, a variance inflation factor (VIF) is applied test to examine multicollinearity across variables (Wooldridge, 2016). If there is a perfect linear relationship between dependent variables, the estimates for a regression are not uniquely computed. Collinearity occurs when two variables have a near perfect linear relationship. If more than two variables are involved it is considered multicollinearity. The main concern of higher degrees of multicollinearity is that the coefficient estimates become unstable and the standard errors may be excessively inflated. Although some portfolio variables, such as HML and RMW, are highly correlated, there were not identified any problems with multicollinearity through the test (Appendix, 12.6). As a rule of thumb, a variable whose VIF figure is greater than 10, should be further examined as it may have multicollinearity issues. All variables have substantially lower values.

Although it is not required for the OLS estimator to be the best linear estimator without bias (BLUE), normally distributed residuals will make it possible to calculate exact t and F values (Wooldridge, 2016). Whether the residuals are normally distributed was tested with a Shapiro-Wilk test (Shapiro and Wilk, 1965). Some of the portfolio regressions appear to have residuals which are not normally distributed. After visually examining the residuals, they are deemed to be acceptable, but certain values do stand out, which probably substantially affect the formal test. This is not unexpected, because when handling large datasets, small deviations will cause a formal test to reject the null hypothesis that the distribution is normally distributed. Upon visual inspection it is concluded that the distribution could have been improved, but that is not entirely unreasonable (Appendix, 12.3 and 12.4). However, in a large sample, the central boundary theorem states that the estimated coefficients distribution moves towards a normal distribution, which implies that normally distributed residuals are not essential (Greene, 2012). What qualifies for a large enough selection to use the central boundary theorem is situation dependent, however, 30 is often used as a rule of thumb. With 192 observations in all of the time series regressions, and with residuals that look relatively normal, it is assumed to not be a substantial problem for the analysis' results.

5.4 Key Takeaways

Methodology
▪ The analysis will utilise portfolios, as they provide less uncertain factor exposure. Employing additional methods, such as panel-based data, would increase the results' robustness. This was, however, not possible due to limited data access at CBS. ▪ The main analysis will study value-based portfolios of 20% of the total selection.   o Various cut-off levels will be studied through robustness tests ▪ Tests of assumption of multiple regression revealed that the datasets had few problems and are regarded as satisfactory for conducting regression analysis.

6 Empirical Analysis

The empirical analysis is divided into several sections. The first section comprises the developed portfolios based on ESG and Controversy scores. These portfolios are analyzed using various pricing models. The classic CAPM is introduced, before applying various multi-factor models. In order to assess the models' results, several robustness tests are applied in the second section, which contribute to greater insights to the findings. In the last section, the results are discussed.

6.1 Portfolio Analysis

6.1.1 Descriptive Statistics

Descriptive statistics for the constructed ESG and Controversy portfolios are illustrated in Table 5. *ESG High* is a value-weighted portfolio that constitutes 20% of the total selection value and comprises companies with the highest ESG scores. *ESG Low* is a value-weighted portfolio that constitutes 20% of the total selection value and contains the companies with the lowest ESG scores. *Controversy High* and *Controversy Low* are divided by the same method but are sorted after controversy scores.

The *ESG High* portfolio has an average of 42 companies per year, while *ESG Low* has 401 companies. *Controversy High* has an average of 259 companies in the portfolio, while *Controversy Low* has an average of 44 companies. Figure 6 depicts the distribution of companies in their respective portfolios for the investigated time period. The correlations between size and score also applies to companies that are not included in the portfolios with the highest and lowest scores, i.e. the companies positioned in the middle three portfolios. This could indicate that there are significant size differences in terms of market value for the companies with low and high scores. It seems that large companies have high ESG scores but low controversy scores. This may also indicate that there exists a correlation between having a high ESG score and low controversy score, which is consistent with the findings made by Blackrock (2018).

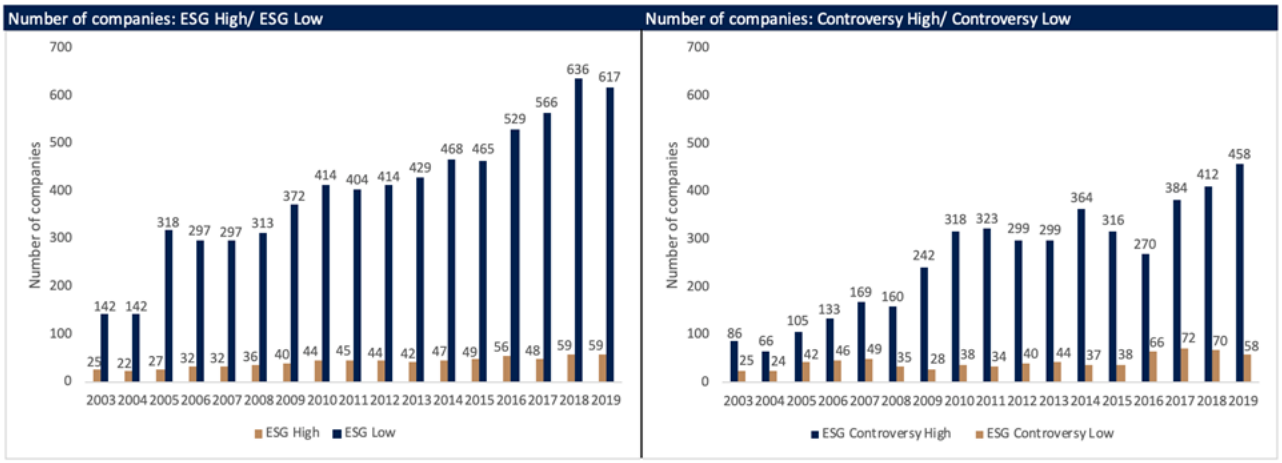


Figure 6: Distribution of companies in portfolios, 2003-2019.

Table 5 shows that the market has a mean return of 1,135% per month. ESG High has a mean return of 0,886% per month (below the market), while ESG Low has a mean return of 1,353% (above the market). The companies with high ESG Scores seem to have lower mean returns than the market, but this may be due to factors other than ESG. For the portfolios sorted by *Controversy High* and *Controversy Low*, the mean return is below the mean return of the market. This may indicate that neither the high nor low controversy portfolios yields the highest returns, but rather the portfolios in-between.

	Mean	Median	Std. dev	Minimum	Maximum
ESG High	0,886	1,553	4,466	-12,939	22,188
ESG Low	1,353	1,978	4,420	-17,392	20,211
Total Market	1,135	1,606	3,977	-12,374	18,074
Controversy High	1,093	1,337	4,162	-15,525	23,151
Controversy Low	0,960	1,355	4,003	-13,384	16,116
RF	0,104	0,035	0,131	0,000	0,440
SMB	0,201	0,190	1,746	-4,640	4,690
HML	0,027	0,055	2,141	-4,980	7,520
WML	0,764	0,945	3,512	-26,100	10,120
RMW	0,324	0,350	1,473	-4,730	4,100
CMA	0,062	-0,020	1,307	-3,530	5,440
ESG factor	-0,466	-0,545	1,779	-5,417	6,621
Controversy factor	0,280	0,289	1,509	-5,887	6,185

Table 5: Descriptive statistics for return and variables. Different portfolios in the first section, variables used for the pricing models in the second section, and self-constructed factors in the third section.

A noteworthy observation is that the HML factor is, on average, 0,027%, i.e. close to zero. The reason behind this result, is that the factor has generally been low since the financial crisis back in 2008. For the self-constructed factors, the ESG factor is negative, that is, companies with high ESG scores tends to have lower mean return than companies with low ESG scores. This is as expected considering the development of *ESG High* and *ESG Low*, which is illustrated in Figure 7. The self-constructed controversy factor is positive and implies that companies with high controversy score have higher mean return than the companies with low controversy score. Both the ESG factor and the controversy factor are aligned with the impression given from the portfolios.

The portfolio with highest standard deviation is *ESG High*, while the portfolio with the lowest standard deviation is *Controversy Low*, illustrated in table 5. All of the portfolios have a standard deviation above the market, which can imply that the portfolios contain unsystematic risk that can be reduced by diversification.

Table 6 illustrates the correlation between the variables that are used in the portfolio analysis. As explained, the ESG factor is strongly correlated with the size factor SMB. In addition, both the ESG factor and controversy factor are negatively correlated with the market. As a rule of thumb, correlation coefficients above 0.70 may imply problems with collinearity. HML and RMW have correlation coefficient of -0,801, which have been addressed and inspected through collinearity tests, as discussed in section 5.3.

	Market	*SMB*	*HML*	*RMW*	*CMA*	*WML*	*ESG factor*	*Controversy factor*
Market	1,000							
SMB	-0,089	1,000						
HML	0,415	0,007	1,000					
RMW	-0,290	-0,088	-0,801	1,000				
CMA	-0,276	-0,154	0,332	-0,341	1,000			
WML	-0,425	-0,003	-0,466	0,342	0,158	1,000		
ESG factor	-0,035	-0,513	0,257	-0,116	0,384	-0,118	1,000	
Controversy factor	-0,263	0,254	-0,147	0,027	0,176	0,306	-0,230	1,000

Table 6 Correlation between different variables used in the analysis

Figure 7 illustrates the indexed return for the four portfolios and the market from the end of July 2003 to the end of June 2019. The portfolios and the market experienced a downturn during the financial crisis in 2008 but have regained a steady growth in the following years. The portfolio for *ESG Low*

has the highest growth, well above the market over time. In the following sections, the relationships will be examined in detail in order to explain the differences in return.

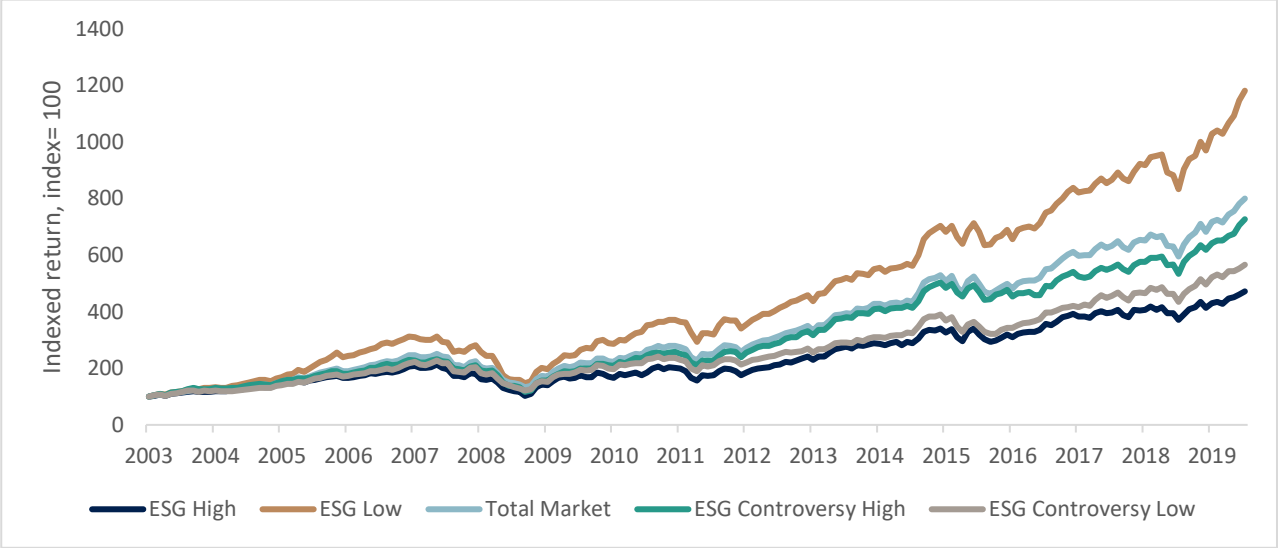


Figure 7: Indexed return for developed portfolios and the market

6.1.2 Sorted for ESG Score

Table 7 contains regressions for the portfolios *ESG High* and *ESG Low* with different explanatory variables. Alpha is significantly negative on a 1 percent level for *ESG High* when using CAPM. This result indicates that portfolios with a high ESG score achieve a lower risk-adjusted return than the market portfolio. When additional risk factors are included, the risk-adjusted returns move towards zero, but are still significantly negative on a 1 percent level for model (2), (3), (4) and (6). Alpha moves toward zero as increasingly more of the portfolios' returns are explained as additional risk factors are introduced. The models suggest that, for the European selection, abnormal returns cannot be achieved by weighting portfolios toward companies with higher ESG scores, given a constant exposure to the remaining risk factors.

Furthermore, the *ESG High* portfolio is negatively exposed to the controversy factor. Thus, some of the return can then seemingly be obtained by investing in companies with a low controversy score. This may indicate that the *ESG High* portfolio on average consist of companies which are involved in many controversies. *ESG High* is positively exposed to the ESG factor, while *ESG Low* is negatively exposed. This regression has been added to test the results, as the portfolios' characteristics imply opposite exposure to the factor.

When the equivalent regressions on the *ESG Low* portfolio are performed, positive alphas are obtained on five out of six models. The only negative alpha is in model (6). However, neither of the models show significant alpha estimates, and thus provide limited insights. *ESG Low* has significant positive exposure to SMB on the 1 percent level, which is expected as the portfolio largely comprises smaller companies. Moreover, it has equally negative significant exposure to HML for all models, except (5).

CMA is significantly negative on a 1 percent level, while RMW is negative for model (4) and (6), but positive in model (5) when the ESG factor is included. Model (6) includes the controversy factor, and companies with a low ESG score have a positive exposure to this factor. This could indicate that companies with a low ESG score are, on average, less exposed to controversies. This is consistent with previous research, which suggests that smaller companies get less public attention compared to larger ones (Aouadi and Marsat, 2018).

By thoroughly studying the included risk-factors, it is possible to examine what differentiates the portfolios beyond differences in ESG. The size factor SMB is always significant at the 1 percent level. For *ESG High* the exposure is negative, while positive for *ESG Low*. Again, the association between size and ESG score is emphasized. Intuitively, portfolios that have positive loadings on the respective factors, also have underlying characteristics which correspond with the factor loadings. A positive SMB loading imply that the portfolio has a small-cap tilt. *ESG High* produces a negative SMB factor in every regression. As the portfolio mainly consists of large cap companies, these results are as expected.

For ESG High, HML is positive and significant at the 1 percent level. In essence, this contradicts the hypothesis presented by Ziegler, Busch and Hoffmann (2011) that companies with high ESG scores have lower systematic risk. Lower systematic risk indicates lower cost of capital and thus a higher valuation. Companies with a high valuation are expected to be negatively exposed to HML. The positive significant exposure to HML would indicate that *ESG High* is largely comprised of value stocks, which explains that the portfolio's returns are attributable to the value premium.

ESG High is not significantly exposed to RMW, while CMA are only significant on a 1 percent level for model (5). *ESG Low* is significantly exposed to CMA on a 1 percent level. Both *ESG High* and

ESG Low have negative exposure to CMA, which may indicate that both portfolios are represented by an overweight of so-called aggressive companies with a higher degree of investments.

In model (4) for *ESG Low*, the HML factor is no longer significant when the RMW factor is included. This is probably related to the fact that these are negatively correlated, and that RMW factor takes over some of the effect that HML previously had. However, these factors are not so extensively correlated that the regression has problems with multicollinearity, as established by the VIF test (appendix, VIF).

Table 7: This table shows coefficients and standard errors from regressions made from monthly returns for portfolios with high/ low ESG score. The portfolios consist of the 20% highest and lowest companies relative to their ESG score. The portfolios are weighted according to the company's market value. The observations indicate the number of months in the timeseries. Avg. No. of companies is average number of companies per month in the portfolio.

ESG High	(1) CAPM	(2) F&F3	(3) Carhart4	(4) F&F5	(5) F&F5 + ESG	(6) F&F5 + Contr.
Mrkt-RF	1,087***	1,039***	1,019***	1,021***	1,049***	1,016***
	(0,020)	(0,021)	(0,021)	(0,024)	(0,018)	(0,023)
SMB		-0,15671***	-0,161***	-0,166***	0,070	-0,123***
		(0,042)	(0,042)	(0,045)	(0,039)	(0,045)
HML		0,191***	0,146***	0,266***	0,104**	0,227***
		(0,0381)	(0,038)	(0,063	(0,050)	(0,062)
WML			-0,078***			
			(0,024)			
RMW				0,078	-0,039	0,054
				(0,0858)	(0,066)	(0,083)
CMA				-0,0948	-0,234***	-0,043
				(0,071)	(0,056)	(0,071)
ESG Factor					0,485***	
					(0,041)	
Controversy Factor						-0,177***
						(0,052)
Alpha	-0,339***	-0,262***	-0,181***	-0,264***	-0,063	-0,212**
	(0,083)	(0,078)	(0,080)	(0,085)	(0,067)	(0,085)
Observations	192	192	192	192	192	192
Avg. No. of companies	42	42	42	42	42	42
Adjusted R^2	0,816	0,825	0,828	0,826	0,848	0,829
ESG Low	(1) CAPM	(2) F&F3	(3) Carhart4	(4) F&F5	(5) F&F5 + ESG	(6) F&F5 + Contr.
Mrkt-RF	1,065***	1,111***	1,111***	1,060***	1,032***	1,064***
	(0,022)	(0,019)	(0,020)	(0,021)	(0,014)	(0,021)
SMB		0,447***	0,447***	0,398***	0,152***	0,369***
		(0,039)	(0,039)	(0,039)	(0,030)	(0,041)
HML		-0,127***	-0,126***	-0,053	0,116***	-0,027
		(0,035)	(0,038)	(0,056)	(0,038)	(0,056)
WML			0,003			
			(0,023)			
RMW				-0,048	0,074	-0,032
				(0,076)	(0,050)	(0,075)
CMA				-0,299***	-0,155***	-0,335***
				(0,063)	(0,042)	(0,064)
ESG Factor					-0,505***	
					(0,031)	
Controversy Factor						0,121**
						(0,047)
Alpha	0,151	0,018	0,0141	0,111	-0,098	0,076
	(0,092)	(0,072)	(0,076)	(0,076)	(0,051)	(0,076)
Observations	192	192	192	192	192	192
Avg. No. of companies	405	405	405	405	405	405
Adjusted R^2	0,799	0,832	0,832	0,837	0,861	0,838

Standard error in parentheses. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$

6.1.3 Sorted for ESG Controversy Score

Table 8 depicts regressions on portfolios with different controversy scores. Considering *Controversy High*, neither of the models generate an alpha significantly different from zero. Hence, there are essentially no evidence that one can achieve higher risk-adjusted excess return by weighting a portfolio toward or away from companies that are subject to few controversies. Moreover, neither of the alphas are statistically significant. Companies with a high controversy score are, on average, negatively exposed to the ESG Factor. This may indicate that companies that are subject to few controversies have lower ESG Scores. This is supportive of previous findings on high/ low ESG score. *Controversy High* has positive exposure to the controversy factor for all models. This is as expected, as the portfolio contains high- performers in this category.

Controversy Low has a non-significant negative alpha when CAPM is applied. As more risk-factors are included the alpha remains negative, but gradually moves toward zero in model (5). Thus, it seems that the portfolio's return can be explained by the exposure to the known risk factors. Companies with a low controversy score are, on average, positively exposed to the ESG Factor, which could imply that companies with many controversies have a higher ESG Score. This is consistent with the rationale presented in table 7, where larger companies tend to have higher ESG scores and get more attention from the overall society.

By studying the risk factors which actually help explain the return, it is shown that it is mainly due to the market exposure. For *Controversy High*, only HML in model (2), WML in model (3) and the constructed controversy factors which produce significant coefficients. A positive exposure to HML is consistent with established research, as small cap is expected to be positively exposed to the factor. The negative exposure to WML can be explained by the portfolio's underperformance compared to the overall market.

For *Controversy Low*, there are several factors that are significant. SMB are significantly negative on a 1 percent level. A negative exposure against SMB could indicate that the portfolio consists of mainly bigger companies. The fact that bigger companies are more exposed to controversies is reasonable, as one could expect that more media attention are drawn to them.

Table 8: This table shows coefficients and standard errors from regressions made from monthly returns for portfolios with high / low controversy Score. The portfolios consist of the 20% highest and lowest companies relative to their controversy score. The portfolios are weighted relative company market value. The observations indicate the number of months in the timeseries. Avg. No. of companies is average number of companies per month in the portfolio.

Controversy High	(1) CAPM	(2) F&F3	(3) Carhart4	(4) F&F5	(5) F&F5 + ESG	(6) F&F5 + Contr.
Mrkt-RF	1,009***	0,994***	0,977***	0,994***	0,989***	1,00***
	(0,020)	(0,022)	(0,022)	(0,025)	(0,025)	(0,023)
SMB		0,076	0,073	0,068	0,028	-0,002
		(0,045)	(0,044)	(0,047)	(0,054)	(0,0455)
HML		0,081**	0,041	0,028	0,056	0,092
		(0,039)	(0,042)	(0,067)	(0,069)	(0,063)
WML			-0,069***			
			(0,026)			
RMW				-0,099	-0,080	-0,061
				(0,090)	(0,091)	(0,084)
CMA				-0,010	0,014	-0,097
				(0,075)	(0,077)	(0,072)
ESG Factor					-0,083	
					(0,057)	
Controversy Factor						0,294***
						(0,053)
Alpha	-0,051	-0,053	0,019	-0,018	-0,052	-0,103
	(0,081)	(0,081)	(0,085)	(0,090)	(0,093)	(0,085)
Observations	192	192	192	192	192	192
Avg. No. of companies	259	259	259	259	259	259
Adjusted R^2	0,809	0,811	0,813	0,811	0,812	0,820
Controversy Low	(1) CAPM	(2) F&F3	(3) Carhart4	(4) F&F5	(5) F&F5 + ESG	(6) F&F5 + Contr.
Mrkt-RF	0,969***	0,947***	0,947***	0,955***	0,963***	0,944***
	(0,019)	(0,020)	(0,021)	(0,023)	(0,023)	(0,019)
SMB		-0,195***	-0,195***	-0,196***	-0,132***	-0,098***
		(0,042)	(0,042)	(0,043)	(0,050)	(0,037)
HML		0,067	0,067	-0,007	-0,051	-0,096
		(0,037)	(0,040)	(0,062)	(0,063)	(0,052)
WML			-0,000			
			(0,0243)			
RMW				-0,108	-0,140	-0,163**
				(0,084)	(0,083)	(0,069)
CMA				0,038	0,000	0,160***
				(0,070)	(0,070)	(0,059)
ESG Factor					0,133**	
					(0,052)	
Controversy Factor						-0,412***
						(0,044)
Alpha	-0,143	-0,083	-0,083	-0,057	-0,002	0,063
	(0,079)	(0,075)	(0,080	(0,083)	(0,085)	(0,070)
Observations	192	192	192	192	192	192
Avg. No. of companies	44	44	44	44	44	44
Adjusted R^2	0,809	0,816	0,820	0,816	0,818	0,823

Standard error in parentheses. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$

6.2 Complementary Analyses and Robustness Tests

This chapter will illustrate three different approaches to the portfolio analysis in order to test the robustness of the results from above. First, the portfolios are divided into 50% and 10% of the market for both *ESG High* and *ESG Low*, as well as *Controversy High* and *Controversy Low*. Secondly, portfolios with an equal number of companies, rather than equal value, are constructed. Lastly, two different time periods are evaluated, 2003-2009 and 2009-2019, in order to illustrate if the relationship between ESG and financial performance is similar over time. The coherence between the time periods could yield some valuable insights from before and after the financial crisis concerning ESG initiatives.

6.2.1 Different Portfolio Sizes

Figure 8 shows the indexed return for portfolios divided into *ESG High* and *ESG Low* with different weightings. The portfolio that consists of the lowest ESG score across the selection, *ESG Low 50%*, have a higher indexed return than the market and far above the companies with the highest ESG scores. This corresponds with the illustration from Figure 7, where *ESG Low* have the highest indexed return, well above the market. Another interesting fact is that *ESG Low 10%* also performs better than the market. This could imply that there exists a linear relationship between the ESG score and the return. Similar to the ESG portfolios that consists of low ESG scores, there seems to be a linear relationship between the portfolios that consists of high ESG scores. Both *ESG High 10%* and *ESG High 50%* performs well below the market. When analyzing portfolios with different weightings of companies with high and low ESG scores, it seems that companies with low ESG scores tend to outperform companies with high ESG scores.

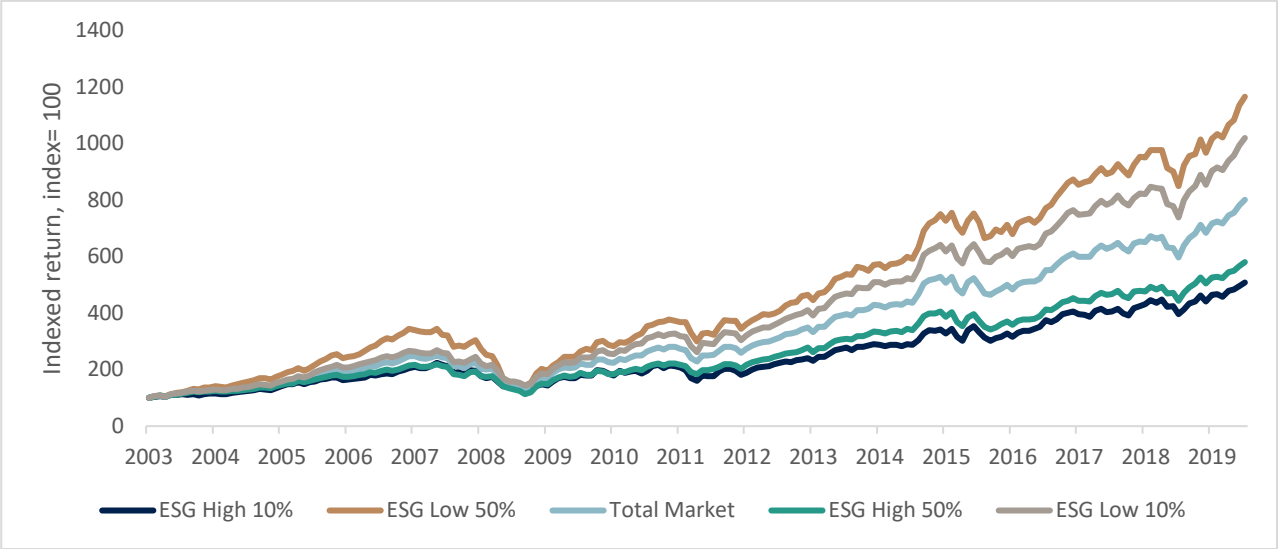


Figure 8: Indexed return for different ESG-portfolios and the market

In Figure 9, the portfolios consisting of high controversy scores looks similar to the portfolios with low ESG scores from Figure 8. The same relationship is evident for the portfolios with low controversy scores and low ESG scores. The portfolio that consists of companies with the lowest concentration of controversies, *Controversy High 50%*, is performing well above the market, similar to *ESG Low 50%* in Figure 8. The portfolio regarding *Controversy High 10%*, is also one of the top performers with approximately the same return as the market. In contrast to the ESG portfolios, it appears to be lower spread between the controversy portfolios up until 2013.

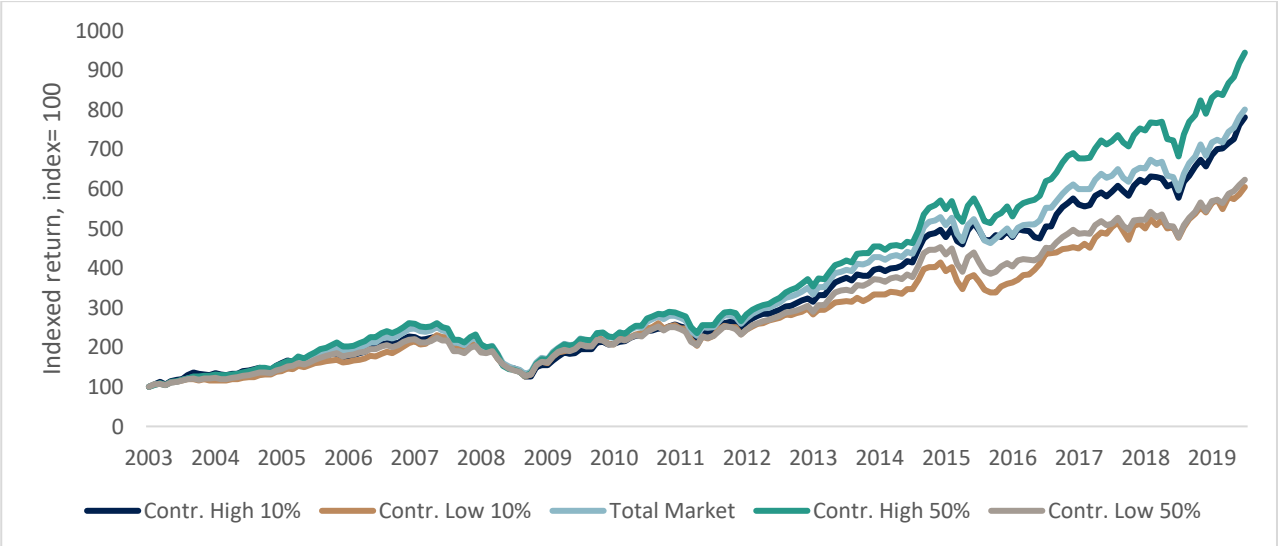


Figure 9: Indexed return for different Controversy-portfolios and the market

Table 9 depicts exposure to the known risk-factors. *ESG High 10%* consists, on average, of 19 companies, while *ESG Low 10%* consists, on average, of 252 companies. The similar differences in the size of the portfolios is found for *Controversy Low 10%* and *Controversy High 10%*. The fact that some of the portfolios consists of only a small number of companies makes it reasonable to assume that some coincidences could have an impact of the results, due to reduced diversification. The results should therefore be interpreted with caution.

ESG High 50% and *ESG Low 50%* have a negative and a positive alpha respectively, while both *ESG High 10%* and *ESG Low 10%* have negative alphas. All of the factors are significant on a 10 percent or 5 percent level for at least one of the models, except for RMW. None of the alphas are significant on any percent level, except model (4). That result could indicate that the companies with the absolute lowest ESG score have a somewhat lower risk-adjusted return. In this case *ESG Low 10%* is merely negative, and close to zero. The result is after the known risk-factors have been taking into account.

Similar to the ESG portfolios, all of the factors for the controversy portfolios are significant for at least one of the models except for the RMW. The alphas are also non-significant. There does not seem to be any excess return from systematically tilting investment towards high/low controversy scores. Both sections show that companies with many controversies appear to be large, and opposite for companies with few controversies.

Table 9: This table shows coefficients and standard errors from various regressions made with Fama & French 5-factor and Carhart momentumfactor on monthly returns for portfolios with high and low ESG scores and high and low controversy scores. The portfolios consist of the 10% best (highest) and 50% lowest (worst) companies, and vice versa, relative to their ESG score/ controversy score. The portfolios are weighted according to the company's market value. The observations indicate the number of months in the timeseries. Avg. No. of companies is average number of companies per month in the portfolio.

ESG	(1)	(2)	(3)	(4)
	ESG High 50%	*ESG Low 50%*	*ESG High 10%*	*ESG Low 10%*
Mrkt-RF	0,958***	1,041***	1,003***	1,100***
	(0,013)	(0,013)	(0,035)	(0,028)
SMB	-0,170***	0,164***	-0,240***	0,530***
	(0,024)	(0,024)	(0,064)	(0,052)
HML	0,091**	-0,106***	0,005	0,031
	(0,036)	(0,036)	(0,096)	(0,078)
WML	-0,038***	0,035**	0,023	0,022
	(0,014)	(0,014)	(0,037)	(0,030)
RMW	0,069	-0,092	-0,169	0,077
	(0,046)	(0,046)	(0,124)	(0,100)
CMA	0,034	-0,040	-0,003	-0,342***
	(0,040)	(0,040)	(0,107)	(0,087)
Alpha	-0,085	0,061	-0,134	-0,009*
	(0,047)	(0,047)	(0,127)	(0,103)
Observations	192	192	192	192
Avg. No. of companies	152	650	19	252
Adjusted R^2	0,861	0,861	0,762	0,815
Controversy	(1)	(2)	(3)	(4)
	Contr. High 50%	*Contr. Low 50%*	*Contr. High 10%*	*Contr. Low 10%*
Mrkt-RF	1,024***	0,975***	0,985***	0,952***
	(0,012)	(0,012)	(0,030)	(0,035)
SMB	0,143***	-0,149***	0,111**	-0,226***
	(0,022)	(0,022)	(0,056)	(0,066)
HML	-0,017	0,003	-0,218***	-0,050
	(0,033)	(0,033)	(0,084)	(0,099)
WML	-0,001	-0,001	-0,006	-0,016
	(0,013)	(0,013)	(0,032)	(0,038)
RMW	-0,026	0,004	0,077	-0,203
	(0,042)	(0,042)	(0,107)	(0,126)
CMA	-0,070	0,065	0,353***	0,022
	(0,037)	(0,037)	(0,093)	(0,110)
Alpha	0,051	-0,079	-0,063	0,038
	(0,043)	(0,043)	(0,110)	(0,130)
Observations	192	192	192	192
Avg. No. of companies	612	183	140	20
Adjusted R^2	0,865	0,864	0,764	0,866

Standard errors in parentheses. *p < 0.10, **p < 0.05, ***p < 0.01

6.2.2 Different Sub-periods

Two separate subperiods have been used to test the robustness of the findings, as well as the impact of specific macroeconomic conditions during the two periods. The use of sub-periods is consistent with studies made by Cohen et al. (1997), Derwall et al., (2005), Kempf and Osthoff (2007), Bollen et al. (2007) and Halbritter et al. (2015). The indexed development in returns in the first subperiod from 2003-2009 are presented in Figure 10. This period is characterized by an upswing in the market which precedes the financial crisis, which seriously hit the market in 2008 and ended in the beginning of 2009, with the subsequent recession. The financial crisis is included in the first subperiod to establish a sense of how the portfolios behave in the economic environment leading up to it, and the effect from it. The crisis was a period characterized by high stock market volatility. This may contribute to some noise in the period, which must be considered when evaluating the validity of the results. Moreover, the first subperiod is substantially shorter than the last which may also contribute negatively to the results. These two periods were chosen rather than a 50/50 split, as it is interesting to study the distinctive differences in macroeconomic conditions in the two periods. However, the regressions were also performed on equally long subperiods, without getting substantially different results in terms of significance.

The period after the financial crisis is characterized by a world economy in recovery, and a record long bull market. For both time periods, the *ESG Low* portfolio is superior to the other portfolios, while *ESG High* is among the lowest performers.

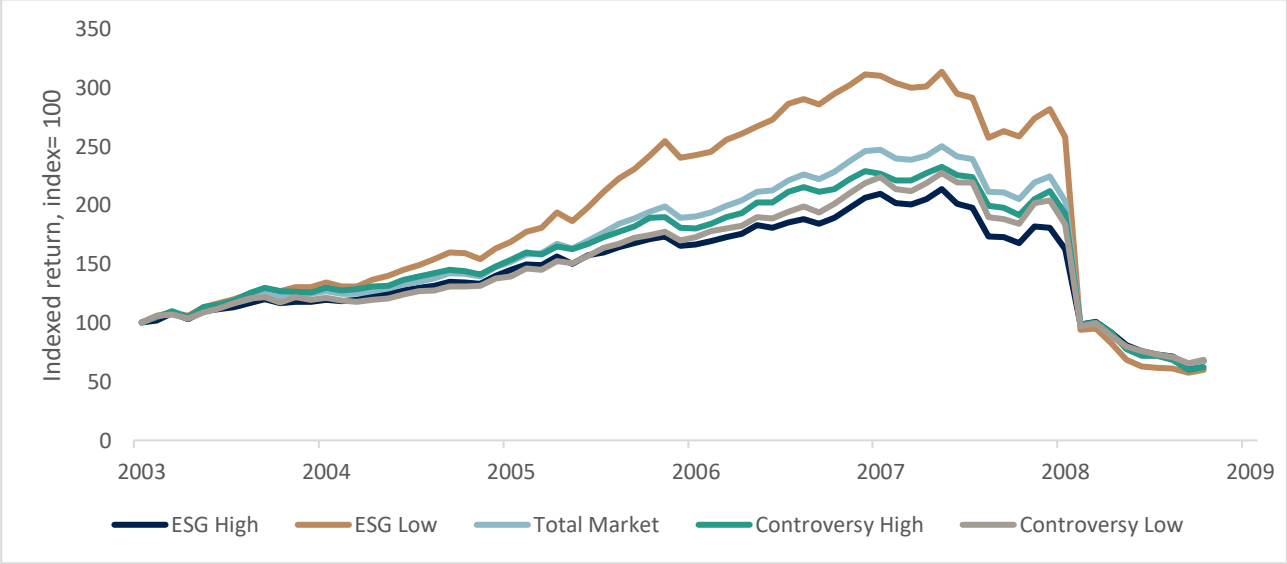


Figure 10: Subperiod 1, 2003-2009

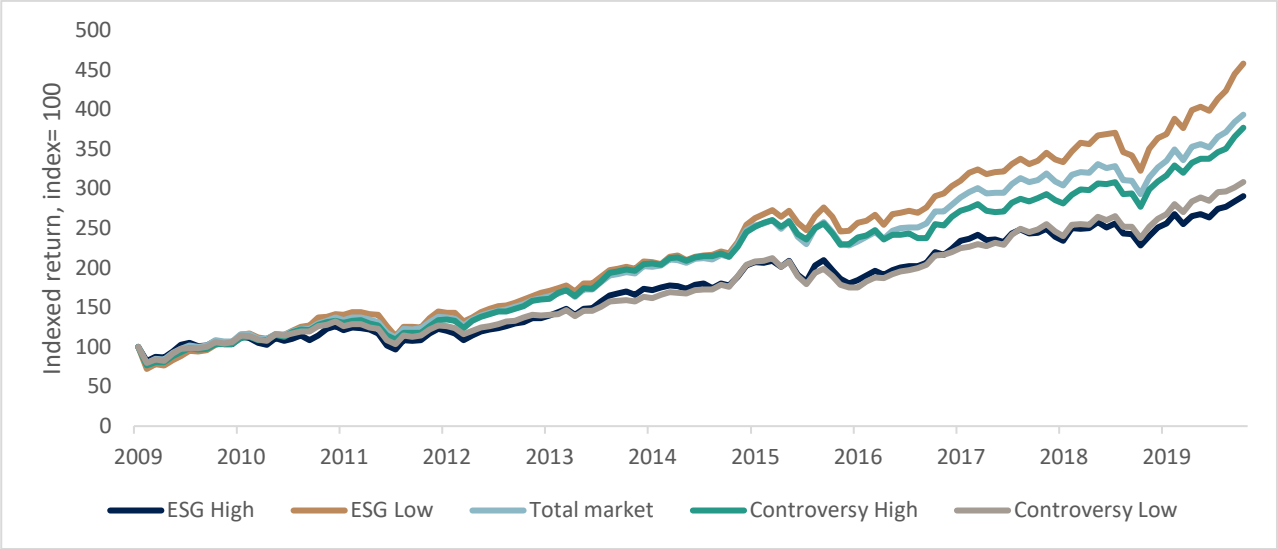


Figure 11: Subperiod 2, 2009-2019

Table 10 illustrates the results from the regressions when divided into two time periods. The portfolios are constructed in the same way as in section 6.1.2 and 6.1.3 for ESG scores and controversy scores, where they were divided into portfolios consisting of the 20% best (highest) and worst (lowest) companies according to their scores. *Period 1* accounts for the period between end of July 2003 until the end of June 2009 and *Period 2* accounts for the period between end of July 2009 until the end of June 2019.

The portfolio for *ESG High* in Period 1 has a significant negative alpha on a 1 percent level. The reason that the alpha is no longer significant in Period 2 may indicate that a possible return difference has been priced in.

Table 10: This table shows coefficients and standard errors from various regressions made with Fama & French 5-factor and Carhart momentum factor on monthly returns for portfolios with high and low ESG scores and high and low controversy scores. The portfolios are divided into two different time periods, Period 1 *(2003-2009) and* Period 2 *(2009-2019. The portfolios consist of the 20% best (highest) and worst (lowest) companies relative to their ESG score/controversy score. The portfolios are weighted according to the company's market value. The observations indicate the number of months in the timeseries. Avg. No. of companies is average number of companies per month in the portfolio.*

Sub-periods, ESG	(1)	(2)	(3)	(4)
	Period 1	Period 2	Period 1	Period 2
	ESG High	*ESG High*	*ESG Low*	*ESG Low*
Mrkt-RF	1,003***	1,018***	1,040***	1,078***
	(0,042)	(0,029)	(0,041)	(0,026)
SMB	-0,077	-0,221***	0,354***	0,433***
	(0,067)	(0,056)	(0,065)	(0,051)
HML	0,300**	0,167**	-0,058	-0,003
	(0,142)	(0,076)	(0,137)	(0,069)
WML	-0,057	-0,061	0,043	0,027
	(0,047)	(0,031)	(0,046)	(0,028)
RMW	0,527***	-0,125	-0,332**	0,078
	(0,152)	(0,096)	(0,148)	(0,087)
CMA	0,77	-0,152	-0,411***	-0,274***
	(0,104)	(0,100)	(0,101)	(0,091)
Alpha	-0,475***	-0,093	0,277	-0,034
	(0,154)	(0,105)	(0,150)	(0,096)
Observations	69	123	69	123
Avg. No. of companies	30	47	261	469
Adjusted R^2	0,824	0,835	0,839	0,836
Sub-periods, Controversy	(1)	(2)	(3)	(4)
	Period 1	Period 2	Period 1	Period 2
	Contr. High	*Contr. High*	*Contr. Low*	*Contr. Low*
Mrkt-RF	0,979***	0,965***	1,037***	0,931***
	(0,042)	(0,029)	(0,048)	(0,027)
SMB	-0,065	0,174***	-0,197**	-0,204***
	(0,067)	(0,056)	(0,076)	(0,053)
HML	0,241	0,019	-0,111	-0,031
	(0,143)	(0,076)	(0,161)	(0,072)
WML	-0,039	-0,082***	-0,069	0,013
	0,048	(0,031)	(0,054)	(0,029)
RMW	-0,340**	0,073	0,163	-0,239***
	(0,153)	(0,096)	(0,173)	(0,090)
CMA	0,121	-0,030	0,131	0,031
	(0,105)	(0,100)	(0,118)	(0,094)
Alpha	-0,162	0,066	-0,041	0,013
	(0,155)	0,105	(0,175)	(0,099)
Observations	69	123	69	123
Avg. No. of companies	120	320	37	47
Adjusted R^2	0,827	0,821	0,805	0,825

Standard errors in parentheses. *p < 0.10, **p < 0.05, ***p < 0.01

6.2.3 Portfolios with Equal Number of Companies

Table 11 illustrates regressions on portfolios that is divided into portfolios with an equal number of companies. These portfolios are sorted after high and low ESG score and high and low controversy score. Every portfolio consists of 20% of the companies that are included for the respective year. There exists, on average, 160 companies in both *ESG High* and *ESG Low* for every year in a portfolio. For the controversies, there exists, on average, 159 companies in both *Controversy High* and *Controversy Low* for every year in a portfolio.

The regression results show that neither of the models have a significant alpha, where the alpha values for *ESG High* and *Controversy Low* are close to zero. However, the alpha values are notably higher for *ESG Low* and *Controversy High*. These findings correspond to the results from the initial analysis, where the portfolios that consists of companies with low ESG scores and companies with high controversy score performs above the market. The regression results show that SMB is significant on 3 out of 4 models, where *ESG High* produces a negative SMB factor which is evident in every regression. An additional significant factor at the 1 percent level is the HML factor. The positive significant exposure to HML indicates that *ESG High* is comprised of value stocks and that the return is attributable to the value premium. The results from this regression, i.e. when portfolios are made with an equal number of companies, does not seem to yield any new insights. The portfolios based on market value seems to yield better results for the analysis. This further emphasizes the clear size relationship described in section 6.1.2, as value-based portfolios seem to include a more homogenous sample of companies, which contribute to more consistent results.

Table 11: The table shows coefficients and standard errors from various regressions made with Fama & French 5-factor and Carhart momentum factor on monthly returns for portfolios with high and low ESG scores and high and low controversy scores. The portfolios consist of the 20% best (highest) and worst (lowest) companies relative to their ESG score/controversy score. The portfolios are constructed with an equal number of companies in each portfolio. The observations indicate the number of months in the timeseries. Avg. No. of companies is average number of companies per month in the portfolio.

Equal number of companies	(1) *ESG High*	(2) *ESG Low*	(3) *Contr. High*	(4) *Contr. Low*
Mrkt-RF	0,967***	1,091***	0,953***	0,966***
	(0,013)	(0,034)	(0,030)	(0,013)
SMB	-0,150***	0,538***	0,046	-0,162***
	(0,025)	(0,063)	(0,055)	(0,024)
HML	0,095**	0,155	-0,082	0,013
	(0,037)	(0,094)	(0,083)	(0,036)
WML	-0,024	0,056	-0,061	-0,009
	(0,015)	(0,037)	(0,032)	(0,014)
RMW	0,067	0,162	0,050	0,001
	(0,048)	(0,121)	(0,106)	(0,047)
CMA	0,017	-0,383	0,049	0,073
	(0,042)	(0,105)	(0,092)	(0,040)
Alpha	0,033	0,116	0,162	0,005
	(0,049)	(0,125)	(0,109)	(0,048)
Observations	192	192	192	192
Avg. No. of companies	160	160	159	159
Adjusted R^2	0,859	0,785	0,776	0,858

6.3 Summary Empirical Analysis

The initial analysis shows that companies with low ESG scores appears to outperform companies with high ESG scores over time. However, when controlling for known risk-factors, the return for *ESG Low* appear to be explained by factors other than ESG. This trend seems to persist throughout the observation period. The exception is *ESG High*, which produces multiple significant alphas. Nevertheless, the robustness test that includes fewer companies in the portfolios shows that companies with the lowest ESG score appear to have a negative alpha, hence, a negative risk-adjusted return.

Assuming that ESG can be considered a risk factor, companies with a high ESG score are associated with lower systematic risk, while companies with a low ESG score are associated with higher systematic risk. If the risk was already priced in, one should expect a portfolio constructed with

companies that consists of low ESG score to have higher expected returns as compensation for increased risk. Ultimately, portfolios constructed with companies with a high ESG score would have a lower expected return. In the first period, this is consistent with the findings for the portfolio with companies with a high ESG score. However, if this was due to a deliberate pricing of risk in the market, it is unexpected that this effect has disappeared in the second period. The analysis did not show any opposite effect for companies with a low ESG score. A possible conscience premium may explain less returns, but it's interesting that the effect is not present in the second period, as one should expect that an increasing number of investors would want to invest in responsible companies.

Another observation, suggesting that the market does not price ESG as a risk factor, is that the companies with the lowest ESG score have a negative alpha. If the market considers companies with a low ESG score to be riskier, the analysis should generate a positive alpha. It is difficult to argue that ESG can be considered as a risk factor when the analysis shows that choosing companies with a low ESG score will make the investment less risky. On the other hand, if ESG is a risk factor and the market has failed to price it accordingly, it can have a negative impact on the return on companies with low ESG scores. In this case, a negative alpha could be explained. The negative alpha could therefore imply that ESG scores captures the effect of underlying risk that negatively impacts those with the lowest ESG scores, and that this risk is not priced into the market.

An interesting finding from the analysis is the relationship between the ESG score and the size of the companies. The ESG factor and SMB (Small minus big) have a correlation of -0,513. The fact that these factors correlate to such great extent may be due to the fact that they consist of many of the same companies. This means that the companies used to construct SMB are also used to construct the ESG factor, and for ESG high and ESG low. If these effects are difficult to distinguish, it may be difficult to identify an ESG effect regardless of the size effect.

Throughout the regression analysis, the self-constructed controversy factor has negative exposure for the portfolio with a high ESG score, and a positive exposure for the portfolio with a low ESG score. This seems a little strange, as there is reason to believe that the companies that have a high ESG score avoids controversies. One possible reason may be that the companies with a higher ESG rating is bigger and thus have more attention and more exposure directed to them. Another reason may be that higher expectations are set for those who promote themselves as a responsible company, which make potential controversies extra harmful compared to companies with less responsible operations.

Regardless of the reason for the positive/ negative exposure, the analysis shows that even if the controversy factor is significant on a 1 percent level, the factor does not have enough impact to change the significance level for alpha. The controversy factor appears to be significant as it explains variance that has already been explained by other factors. Consequently, the inclusion of the self-constructed controversy factor does not seem to yield much insights.

For the portfolios constructed on the controversy scores, the initial analysis shows that portfolios with high and low controversy scores have a lower return than the market. In addition, neither of the alphas are significant after accounting for the known risk-factors. After dividing the time period into two sub-periods, the regressions do not show any significant alphas for any of the sub-periods. When the companies are divided into portfolios with an equal number of companies in each, or with a 50% or 10% split, the same results are provided, i.e. a non-significant alpha. There exists no evidence in this analysis that can provide any answers if it's possible to use controversy scores to exploit mispricing in the aftermath of controversies, nor that controversies have any effect on the return. Additionally, there is no evidence that the market doesn't consider controversy related to the controversy score as an indicator of underlying systematic risk. Thus, the market appears quite efficient in terms of controversies.

6.4 Key Takeaways

Empirical Analysis

- The main analysis shows that the ESG Low portfolio clearly outperforms ESG High. The findings are consistent across nearly all portfolio sizes
  - An interesting exception is that ESG Low 10% produces a negative significant alpha, which implies that the poorest ESG performers are not attractive from an investment perspective
- Significant negative alphas are found for ESG High across all models
- The portfolios exposure to the known risk factors is as expected considering the portfolios' characteristics, e.g. ESG High has negative exposure to SMB as the portfolio comprises large cap stocks, and vice versa for ESG Low
- The analysis fails to produce significant alphas for either of the controversy portfolios

7 Discussion

This chapter will discuss the results from the main analyses on both ESG and Controversies in relation to a selection of the existing research presented in the literature review. Research implying both a positive and negative relationship between ESG and financial performance is included, and the selection is based on what results are deemed most comparable. The chapter will further involve the results from the robustness analysis to help shed light on possible relationships.

7.1 ESG

The analysis only shows a consistently significant relationship between ESG scores and financial return when considering the *ESG High* portfolio and the entire period. The relationship remains significant and negative for both CAPM and when additional control variables are introduced in Carhart and Fama & French three- and five factor variations. The interesting exception is model (5) where it is also regressed on the ESG factor, which as discussed, is expected due to the portfolio being regressed against its own characteristics. A notable observations is that the *ESG High* portfolio produces a significant negative alpha even after including the controversy factor in model (6). This indicates that ESG scores contribute to negative excess returns, which are not explained by variations in controversy score. This could imply that its underperformance relative to the market is due to its higher ESG performance. The authors fail to prove similar significant relationships, neither negative nor positive, for ESG Low and the controversy portfolios. The insights from these results will be discussed in light of existing research to investigate whether conclusions may be drawn.

When the analyses are specified in the robustness section, both in smaller and bigger portfolios as well as over different subperiods, a few significant results are obtained. In contrary to previous findings, the analysis find evidence that the companies with the absolute lowest ESG scores have negative excess return. However, the negative alpha is not substantially different from zero. This implies that, relative to the risk associated with the portfolio, the return is inferior to the market benchmark. This indicates that the lowest possible ESG score is not desirable from an investor perspective. Drawing the same comparison for the other 10% and 50% portfolios, more similar relationships across portfolio sizes are indicated for both *ESG High* and *Low*; the results do not differ substantially from the main analysis. These similarities across portfolio sizes and ESG scores increase the robustness of the main findings and reduce the chance that they are a result of coincidence.

The sub-period analysis also shows a statistically significant negative alpha for the ESG high portfolio in the first period, which is no longer significant in the second period. The primary explanation to these differences is the differences in the underlying datasets. Such variations in results are signs that the portfolios are sensitive to differences in number of observations, time and macroeconomic conditions. Reduced data, i.e. shorter time periods and fewer companies in each portfolio, weakens the confidence in the findings. This is because the respective portfolios may experience reduced diversification effect and fewer observations reduces reliability. When the portfolios are constructed, various samples from the total dataset are considered, and observations outside these portfolios are thus not included. Another element that comes into play is the use of value- weighted returns. Both ESG and controversy scores have clear positive correlation with company size. It is reasonable to assume differences in results if the weights had been treated differently, e.g. equally weighted. However, value-weighting portfolio return is common practice, and applying the same method across all analyses secures consistency and comparability of the results.

As discussed in the literature review, Zhang (2017) opines that there are no disadvantages of performing responsible investments. On the contrary, she rather argues that it could possibly add value, but that the associated risk may already be priced in. If ESG risk is priced in, risk compensation should ensure that the portfolios with a high ESG score achieve negative alpha, and the ESG Low portfolios should produce a positive alpha, everything else equal. The results show strong indication that this is the case, as the ESG High portfolio has significant negative alphas and are consistently outperformed by ESG Low. The ESG Low portfolio returns positive alphas regardless of model, but fail to produce significant results. These results are supportive of Zhang's (2017) rationale that the information that ESG scores comprise has been evaluated by the market and is reflected in the stock price. Moreover, the portfolio with the 10% lowest performing companies with respect to ESG has a significant negative alpha. This implies that ESG represents a risk factor which in particular also affect the lowest ESG performers negatively, and that this effect is not priced in. The results imply that, if ESG is associated with risk, a positive correlation between ESG score and return should be seen, given that the risk is not already priced in, and a negative correlation if it is priced in. The results of Zhang (2017) can help explain our own findings. However, the results from the regression analysis stand in stark contrast the research implying a clear positive relationship between ESG and financial performance, such as Russo and Fouts (1997), Hart and Ahuja (1996).

The analysis describes a clear positive correlation between size and ESG score. ESG High has on average 42 companies per year, while ESG Low has 405. The findings are supportive of Bianchi and Noci (1998) and suggests that a proactive ESG strategy is problematic for SMEs as they lack the resources and skills to implement them. Hence, they receive relatively lower scores. The primary analysis also shows that controversies are associated with greater firm value, findings similar to Aouadi and Marsat (2018). The size/score relationship is supported by the results from portfolios with equal number of companies, which develop far less significant results. This implies that developing portfolios on the basis of value creates more homogenous samples with similar exposures to risk factors.

As discussed, ESG High has comparably weaker returns. This could also be explained by the applied restrictions which reduce the investment universe. There is a risk of developing portfolios with an overweight in equities with unattractive risk-adjusted returns to meet predefined ESG standards, as suggested by Climent and Soriano (2011). Moreover, Fama and French (1992) find that small-capitalization stocks outperform large-capitalization stocks throughout their research, which is supportive of the primary results of this thesis. This could imply that the weaker returns are not only explained by the respective ESG scores, but are due to consequences of a limited investment universe and company size.

Oberndorfer et al. (2013) tests the effect of firm performance for corporations in two different sustainability stock indexes by using CAPM and the three-factor model. Their results imply negative average cumulative abnormal returns for both indexes. However, their results are non-significant on DJSI STOXX, while it is significant on DJSI World. The results made from Oberndorfer et al. (2013) is similar to the significantly negative results found for the portfolios with high ESG scores, and could further imply that these results could be due to the market index that is used. However, the authors do not find evidence that systematically investing in only the poorest ESG performers generate statistically significant positive alpha over time.

Interestingly, the findings on ESG seem to contradict those of Revelli and Viviani (2015), which find that there are no real advantages or disadvantages associated with responsible investments. This analysis imply a negative relationship between ESG performance and financial return. However, they too claim that results are highly dependent on method choices and data sources. While their study solely considers ESG, our analysis find comparable relationships for controversy scores. This is

admittedly before taking any costs associated with conducting an analysis of corporate responsibility into account. The findings of Revelli and Viviani (2015) are supportive of those of Halbritter and Dorfleitner (2015), which studied the U.S market from 1991 to 2012. Their portfolio analysis does not state a significant return difference between companies with high and low ESG scores, and neither generated abnormal returns. That the findings in this thesis contrast those on other selections and time periods is an interesting and notable observation.

Moreover, Friede, Bush and Bassen (2015) also found indications of varying relationships in a meta-study combining more than 3700 study results. They refer to multiple previous studies which find neither a positive nor negative correlation. Furthermore, they found that a substantial portion of the findings showed a positive relationship, but these findings where only particularly evident in the North American (65,4%) and emerging markets (42,7%). Looking at developed markets in Europe, they only found a positive correlation in 26,1% of the studies. These results imply a clear difference in how ESG is perceived in the world's major financial markets, with the U.S. clearly being more inclined to view it as a positive contribution. However, the study of European ESG portfolio performance support a majority of findings up until 2015. These findings support the notion that existing research is not consistent on the matter of ESG and financial returns.

7.1.2 ESG, Sub-Conclusion

Currently, there does not exist a consensus among researchers on the relationship between ESG and financial performance. Our results, however, are more supportive of the literature implying a negative correlation between increased ESG initiatives and financial performance measured by excess returns. It should be noted, that the results from the analysis do not propose a realistic investment strategy. It is unlikely that a rational investor would construct a portfolio based solely on companies' ESG performance, as multiple other factors and considerations are deemed as relevant to return. This analysis is however a theoretic study of the explicit relationship between ESG and risk-adjusted returns. However, as discussed in section 1.1, multiple funds have ESG incorporated in their investment mandate. The results from the analysis should thus be viewed relative to a broader investment strategy.

7.2 Controversies

For ESG-related controversies, which is a fairly new subject area, the portfolio analysis doesn't yield any significant results. This outcome accounts for both the aggregate level and on a more specific level. The primary analysis show non-significant negative alphas for 10 out of 12 models. The lack of significance for these return values makes it difficult to draw conclusions on the portfolios' risk-adjusted performance.

As discussed in section 4.5.2, being involved in many controversies, will result in a lower score. Garvey et al. (2016) argues that controversies are associated with negative returns. The statistical analysis does not show clear indications of this, however, the mean return of Controversy Low is weaker than that of Controversy High. This result is indicates some support of the conclusions of Garvey et al. (2016). Aoudi and Marsat (2016) concluded that controversies are associated with higher returns for "high-attention" companies, but argues that this relationship expires when other measures for responsibility are included. Normally, "high-attention" companies would be the larger ones experiencing more scrutiny from stakeholders. As neither of the controversy portfolios return significant results, it is not possible to conclude if the findings are supportive of this argument or not.

Section 2.5 describes multiple intuitions as to how controversies may contribute to excess returns. A common rationale is the mispricing after a controversy occurs, before the market has evaluated the full implications of the controversy. The revaluation would then move towards the company's fundamental value. Given this interpretation, one could expect more significant results in the second sub-period, assuming that the market would be better equipped to not misprice the consequences of a controversy. However, the results do not show any clear implications of this.

Given the clear correlation between number of controversies and company size, it is reasonable to believe that controversies systematically happen more often for certain companies. Further, Blackrock (2018) suggests that controversies are positively correlated with previous controversies.
It could, however, be reasonable to believe that companies which are more transparent about their ESG policies should endure fewer controversies. Despite this, Blackrock (2018) found results arguing against this: Firms that claim to employ the most elaborate ESG policies were involved in the most controversies. They were more exposed to lawsuits and government interactions over controversies related to hiring, tax issues and discrimination. By contrast, companies that disclosed the fewest ESG

policies seem to be involved in far fewer controversies. Blackrock (2018) here emphasizes size as one of the most explanatory factors when it comes to predicting whether a company gets entangled in controversies the very next year, where the biggest companies are most exposed.

It can be argued that investors should be compensated for the risk involved when investing in companies with many controversies, where a negative relationship between controversy score and return can be justified by such a theory of compensation for risk. An important note is that the analysis doesn't find any evidence supportive of Blackrock's (2018) rationale, but neither any results which argue against it. The analysis find no significant effect controversies on excess returns, neither in the initial analysis nor in the robustness tests when portfolios are divided into various sizes and sub-periods. This weakens the credibility of the results and raises concerns on whether there is any noticeable connection between controversies and financial performance or that it is possible to exploit any excess returns due to controversies.

7.3 Key Takeaways

Discussion

- The results show a significant negative relationship between ESG and financial performance for the analysed selection.
- These results are supportive of previous studies conducted on the European market, but differ substantially from what has previously been observed in North-America.
- The results are not to be considered as realistic investment strategies, but rather as an academic study of ESG performance relative to financial performance.
- While ESG High produces significant negative alphas across multiple models, significant alphas are not obtained for ESG Low. This means that, relative to the risk involved, the portfolio does not produce significant abnormal returns.
- *Controversy Low* models find negative alphas, which are supportive of previous findings that suggest controversies have a negative impact on financial returns. However, they are not significant and thus provide limited insights.

8 Weaknesses

When European companies that do not provide information on corporate responsibility are excluded from the dataset, the selection consists of a total of 1,745 companies. Further excluding those with inadequate or missing data on market value of equity and return in Datastream, the number is reduced to 1,254 companies. As a result, a larger number of companies is lost due to the lack of data, and especially due to lack of ESG and controversy data. A significant reduction of the selection could imply a risk that the companies used in the analyses are not representative of the European market. If the selection is unable to reproduce known relationships between risk factors and returns, it raises doubt as to whether the analysis is able to credibly show how ESG and controversies actually affect returns. If this is the case, a more representative selection may have produced substantially different results. The applied selection comprises most of the companies in Europe which Datastream has recorded ESG data on. If the case is that the selection is not representative, it not only casts doubt on the results of this thesis, but also on previous studies on ESG and returns. At least if the studies are based on data retrieved from Datastream. This issue may also be present in studies concerning other markets than the European. Thomson Reuters has a more comprehensive, and possibly more representative, data for North America, but this does not exclude similar problems for American studies. It could be reasonable to assume that possible differences are due to systematic differences between companies that report enough data to construct ESG scores, and those that do not. Another explanation is capacity constraints with respect to gathering data to construct scores.

One of the most prominent sources of uncertainty in the thesis may be the choice of data provider for the ESG and controversy scores. There is a lack of correlation and consistency from different data providers with regards to ESG. As discussed, the ESG data providers use different methods when analyzing responsible actions for companies, where the measurement divergence explains 50% of the overall divergence between the individual providers. A "rater effect" has been exposed for the five most used ESG data providers, which means that the individual categories are influenced by the respective provider's view of the company. This indicates that the providers use subjective measures on the overall score, where they can value e.g. governance to have a higher impact for the overall score than the environmental factor. Ultimately, the individual results in the analysis could differ to a greater or lesser extent if another data provider were chosen.

This thesis focuses solely on portfolio analysis. A panel data analysis would investigate the relationship between ESG and financial performance on company level. When only analyzing portfolios, details in company specific information may be missed. However, previous research has implied less uncertainty in factor exposure for portfolios, but similar standard deviations for risk premium estimates. This is because grouping companies in larger portfolios will diversify much of the information about each company's factor exposure. Hence, analyzing portfolios could imply large efficiency losses compared to company specific analysis. In that regard, it would be interesting to perform both portfolio and panel data analyses and compare the results. Applying more than one method would also increase the robustness of the results. However, due to restricted data access, this was not possible.

The analysis revealed a clear correlation between both ESG and controversy scores and company size. As most of the portfolios are constructed on the basis of value, the high ESG performers comprise substantially fewer companies than the lower performing ones. This could lead to reduced diversification effect in the higher performing portfolios. Moreover, the analysis is based on annual observations for ESG and controversy scores. There may be a substantial time between the month controversy scores are being used, and when the actual controversy became known to the market. This could be somewhat misleading. If the scores were updated more frequently, a clearer picture of the development of returns as a result of controversies could be made.

9 Conclusion

This thesis has examined the effect of ESG and controversy conditions on companies' returns in Europe from 2003-2019. The effect has been investigated by looking at self-constructed portfolios, where the portfolios have been examined on different dimensions and time-periods. As a measurement of a company's responsible actions and initiatives, ESG scores and controversies scores have been collected from Thomson Reuters Datastream. More specifically, the paper examines the effect of various score levels, and not the explicit change in score or whether it has any effect at all to have been assigned a score. The dataset is extensive and consists of 1,254 unique companies in Europe. However, certain companies are still excluded, mostly due to the lack of ESG or controversy scores.

In order to analyze possible relationships, various methods and calculations are applied to handle potential challenges in the datasets. The portfolio construction, data preparation and portfolio analysis follow standards from the literature. In addition, relevant discussions related to the analyses, previous empirical research on the subject and other published articles have been taken into consideration and have been evaluated during the process.

The main findings of the analysis show that there is a statistically significant negative relationship between high ESG scores and the risk-adjusted return. Moreover, the analysis fails to produce statistically significant alphas for the portfolio with high ESG scores in the robustness test. The analysis also found that portfolios with low ESG scores tend to outperform companies with a high ESG score over time. However, the analysis shows that the absolute bottom performers with concerns to ESG, tend to have a negative alpha. The authors detected no evidence in the analysis that can provide any answers if it's possible to use controversy scores to exploit mispricing in the aftermath of controversies, nor that controversies have any effect on the risk-adjusted return. Thus, the market appears quite efficient in terms of controversies. The significant results observed between responsibility and risk-adjusted return are not consistent across the robustness tests and it is therefore not possible to reach a consensus about the relationship. Lack of consistency makes it difficult to argue that ESG and controversies actually has an impact on risk-adjusted return. However, the fact that the analysis found multiple significant results makes it difficult to exclude the fact that there exists, or has existed, some form of correlation.

For further research, the authors would recommend investigating the relationship by using other ESG-data providers. These providers are proved to supply investors with highly dissimilar ESG scores, as they evaluate all three pillars (E,S,G) differently. Another proposal would be to investigate the ESG pillars individually. This could be to focus solely on the environmental factor and see if it has any impact on the financial performance of a company. Further research could include a panel data analysis on the European market, to contribute to the work already performed by Ortas et al. (2015). It may be interesting to investigate how the relationship evolves in other geographical locations, such as Eastern/Western Europe or Scandinavia/UK, or on an industry-level. Another compelling subject area would be to explore the differences between the companies that report on responsible actions and have an ESG score and compare them to companies that do not have a score. This could be to inspect if they appear to have contrasting exposure to the known risk-factors, or if the ESG score has any impact on the attention that the companies receive. The authors also encourage future research to concentrate on longer time-series data or perform a similar analysis in the future, e.g. in 10 years' time. A similar analysis would arguably provide different results, as ESG considerations are becoming increasingly more relevant among consumers, government, investors and stakeholders overall, and would arguably be even more relevant in the years to come.

10 Bibliography

Allen, F., Carletti, E. and Marquez, R. (2007) Stakeholder capitalism, corporate governance and firm value. Univ.-Bibliothek Frankfurt am Main.

Ambec, S. and Lanoie, P. (2007) When and why does it pay to be green?

Ang, A. (2014). Asset Management: A Systematic Approach to Factor Investing. United States: Oxford University Press.

Aouadi, A. and Marsat, S. (2018) 'Do ESG controversies matter for firm value? Evidence from international data', Journal of Business Ethics. Springer, 151(4), pp. 1027–1047.

Aupperle, K. E., Carroll, A. B., and Hatfield, J. D. (1985). An Empirical Examination of the Relationship between Corporate Social Responsibility and Profitability. The Academy of Management Journal, 28(2), 446–463.

Banz, R. W. (1981). The relationship between return and market value of common stocks. Journal of Financial Economics, 9(1), 3–18.

Barnea, A. and Rubin, A. (2010) 'Corporate social responsibility as a conflict between shareholders', Journal of business ethics. Springer, 97(1), pp. 71–86.

Barney, J. B. and Hansen, M. H. (1994) 'Trustworthiness as a source of competitive advantage', Strategic management journal. Wiley Online Library, 15(S1), pp. 175–190.

Baron, D. P. (2001) 'Private politics, corporate social responsibility, and integrated strategy', Journal of Economics & Management Strategy. Wiley Online Library, 10(1), pp. 7–45.

Bassen, A., and Kovàcs, M. (2008). 'Environmental, Social and Governance Key Performance Indicators from a Capital Market Perspective», vol.9, no.2, pp.182-192

Bauer, R., and Hann, D. (2010) 'Corporate Environmental Management and Credit Risk', (European Centre for Corporate Engagement (ECCE)

Bauer, R., Koedijk, K., and Otten, R. (2005). International evidence on ethical mutual fund performance and investment style. Journal of Banking and Finance, 29(7), 1751–1767.

Becchetti, L. and Ciciretti, R. (2009) 'Corporate social responsibility and stock market performance', Applied financial economics. Taylor & Francis, 19(16), pp. 1283–1293.

Bello, Z. Y. (2005) 'Socially responsible investing and portfolio diversification', Journal of Financial Research. Wiley Online Library, 28(1), pp. 41–57.

Berg, F, Kölbel, J and Rigobon, R, (2019). Aggregate Confusion: The Divergence of ESG Ratings. MIT Sloan Research Paper No. 5822-19.

Berman, S. L. et al. (1999) 'Does stakeholder orientation matter? The relationship between stakeholder management models and firm financial performance', Academy of Management journal. Academy of Management Briarcliff Manor, NY 10510, 42(5), pp. 488–506.

Bianchi, R. and Noci, G. (1998) '" Greening" SMEs' Competitiveness', Small Business Economics. Springer, 11(3), pp. 269–281.

Black, F. (1995). Estimating Expected Return. Financial Analysts Journal, 51(1), 168–171.

Blackrock. (2018). Sustainable investing: a "why not" moment Environmental, social and governance investing insights. Retrieved 7 April 2020, from https://www.blackrock.com/corporate/literature/whitepaper/bii-sustainable-investing-may-2018-international.pdf

Blitz, D., X. Hanauer, M., Vidojevic, M., and van Vliet, P. (2018). Five Concerns with the Five-Factor Model. The Journal of Portfolio Management, 44 (4) 71-78

Bloomberg. (2020). Environmental, social and governance (ESG) data Content and Data. Retrieved 22 March 2020, from https://www.bloomberg.com/professional/solution/esg/

Bodie, Z., Kane, A. and Marcus, A. J. (2017). Investments. McGraw-Hill Irwin.

Bollen, N. P. B. (2007) 'Mutual fund attributes and investor behavior', Journal of Financial and Quantitative Analysis. Cambridge University Press, 42(3), pp. 683–708.

Bondarenko, P. (2019). Enron scandal | Summary, History, & Facts. Retrieved 1 April 2020, from https://www.britannica.com/event/Enron-scandal

Bouslah, K., Kryzanowski, L., and M'zali, B. (2013). The impact of the dimensions of social performance on firm risk. Journal of Banking and Finance, 37(4), 1258–1273.

Brekke, K. A. and Nyborg, K. (2004) Moral hazard and moral motivation: Corporate social responsibility as labor market screening. Memorandum.

Breusch, T. S. (1978) 'Testing for autocorrelation in dynamic linear models', *Australian Economic Papers*. Wiley Online Library, 17(31), pp. 334–355.

Breusch, T. S. and Pagan, A. R. (1979) 'A simple test for heteroscedasticity and random coefficient variation', *Econometrica: Journal of the Econometric Society*. JSTOR, pp. 1287–1294.

Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. Journal of Finance, 52(1), 57–82.

Chollet, P., and Sandwidi, B. W. (2018). CSR engagement and financial risk: A virtuous circle? International evidence. Global Finance Journal, 38, 65–81.

Climent, F. and Soriano, P. (2011) 'Green and good? The investment performance of US environmental mutual funds', Journal of Business Ethics. Springer, 103(2), pp. 275–287.

Cochrane, D. and Orcutt, G. H. (1949) 'Application of least squares regression to relationships containing auto-correlated error terms', *Journal of the American statistical association*. Taylor & Francis Group, 44(245), pp. 32–61.

Cohen, M., Fenn, S., and Konar, S. (1997). Environmental and Financial Performance: Are They Related?.

Cook, R. D. and Weisberg, S. (1983) 'Diagnostics for heteroscedasticity in regression', *Biometrika*. Oxford University Press, 70(1), pp. 1–10.

Darnall, N., Henriques, I. and Sadorsky, P. (2010) 'Adopting proactive environmental strategy: The influence of stakeholders and firm size', Journal of management studies. Wiley Online Library, 47(6), pp. 1072–1094.

Della Croce, R., Stewart, F. and Yermo, J. (2011). Promoting Longer-Term Investment by Institutional Investors. OECD Journal: Financial Market Trends, 2011(1), 145-164.

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

Dhaliwal, D., Li, O. Z., Tsang, A., and Yang, Y. G. (2014). Corporate social responsibility disclosure and the cost of equity capital: the roles of stakeholder orientation and financial transparency. Journal of Accounting and Public Policy, Vol. 33, 328-355.

Diltz, J. D. (1995). Does Social Screening Affect Portfolio Performance? The Journal of Investing, 4(1), 64 LP – 69.

Dickey, D. A. and Fuller, W. A. (1979) 'Distribution of the estimators for autoregressive time series with a unit root', *Journal of the American statistical association*. Taylor & Francis, 74(366a), pp. 427–431.

Dorfleitner, G., Halbritter, G., and Nguyen, M. (2014). Measuring the level and risk of corporate responsibility - An empirical comparison of different ESG rating approaches. Universitätsstraße 31, 93053 Regensburg, Germany: University of Regensburg, Department of Finance.

Dowell, G., Hart, S., and Yeung, B. (2000). Do Corporate Global Environmental Standards Create or Destroy Market Value? Management Science, 46(8), 1059–1074.

Durbin, J. and Watson, G. S. (1992) 'Testing for serial correlation in least squares regression. I', in *Breakthroughs in Statistics*. Springer, pp. 237–259.

Eccles, R., and Klimenko, S. (2019). Shareholders Are Getting Serious About Sustainability. Harvard Business Review, from the May-June 2019 Issue. Retrieved 17 April 2020, from https://hbr.org/2019/05/the-investor-revolution

El Ghoul, S., Guedhami, O., Kwok, C. C. Y., and Mishra, D. R. (2011). Does corporate social responsibility affect the cost of capital? Journal of Banking and Finance, 35(9), 2388–2406.

Escrig-Olmedo, E., Fernández-Izquierdo, M. Á., Ferrero-Ferrero, I., Rivera-Lirio, J. M., and Muñoz-Torres, M. J. (2019). Rating the Raters: Evaluating how ESG Rating Agencies Integrate Sustainability Principles. Sustainability, 11(3), 915.

Euribor. (2020). What is Euribor?. Retrieved 4 April 2020, from https://www.euribor-rates.eu/en/what-is-euribor/

Fama, E. F., and French, K. R. (1992). The Cross-Section of Expected Stock Returns. The Journal of Finance, 47(2), 427–465.

Fama, E. F., and 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., and French, K. R. (1996). Multifactor explanations of asset pricing anomalies. Journal of Finance, Vol. 51 (No. 1), 55-84.

Fama, E. F., and French, K. R. (2015). A five-factor asset pricing model. Journal of Financial Economics, 116(1), 1–22.

Fisman, R., Heal, G. and Nair, V. (2006) 'A model of corporate philanthropy', Columbia University and University of Pennsylvania.

Freeman, R. E. and Reed, D. L. (1983) 'Stockholders and stakeholders: A new perspective on corporate governance', California management review. SAGE Publications Sage CA: Los Angeles, CA, 25(3), pp. 88–106.

French, K. (2020). Kenneth R. French - Description of Fama/French Benchmark Factors. https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/f-f_bench_factor.html

French, K. (2020). Kenneth R. French - Data Library. Retrieved 15 April 2020, from http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Friede, G., Busch, T., and Bassen, A. (2015). ESG and financial performance: aggregated evidence from more than 2000 empirical studies. Journal of Sustainable Finance & Investment, 5(4), 210–233.

Friedman, M. (1970). 'Social responsibility of business', New York Times Magazine.

Friedman, M., and Friedman, R. D. (1962). Capitalism and freedom. Chicago London.

Galbreath, J. (2013). 'ESG in Focus: the Australian Evidence', Journal of Business Ethics, vol.118, pp. 529-541

Galema, R., Plantinga, A., and Scholtens, B. (2008). The stocks at stake: Return and risk in socially responsible investment. Journal of Banking and Finance, 32(12), 2646–2654.

Garvey, G. T. et al. (2016) 'A pitfall in ethical investing: ESG disclosures reveal vulnerabilities, not virtues', Not Virtues (September 19, 2016).

Global Sustainable Investement Alliance (GSIA). (2019). 2018 Global Sustainable Investment Review. Retrieved 4 April 2020, from http://www.gsi-alliance.org/wp-content/uploads/2019/06/GSIR_Review2018F.pdf

Greene, W. H. (2012) 'Econometric analysis, 71e', *Stern School of Business, New York University*.

Gregory, A., Tharyan, R., and Whittaker, J. (2014). Corporate Social Responsibility and Firm Value: Disaggregating the Effects on Cash Flow, Risk and Growth. Journal of Business Ethics, 124(4), 633–657.

Griffin, J. J., and Mahon, J. F. (1997). The Corporate Social Performance and Corporate Financial Performance Debate: Twenty-Five Years of Incomparable Research. Business & Society, 36(1), 5–31.

Guba, E., and Lincoln, Y. (1994). Competing paradigms in qualitative research. Handbook of qualitative research, 105-117.

Halbritter, G., and Dorfleitner, G. (2015). The wages of social responsibility — where are they? A critical review of ESG investing. Review of Financial Economics, 26(1), 25–35.

Hall, P., and Padmanabhan, A. R. (1992). On the bootstrap and the trimmed mean. Journal of Multivariate Analysis, 41(1), 132–153.

Harjoto, M., and Jo, H. (2015). Legal vs. Normative CSR: Differential Impact on Analyst Dispersion, Stock Return Volatility, Cost of Capital, and Firm Value. Journal of Business Ethics, 128(1), 1–20.

Hart, S. L., and Ahuja, G. (1996). Does it pay to be green? An empirical examination of the relationship between emission reduction and firm performance. Business Strategy and the Environment, 5(1), 30–37.

Heal, G. (2005) 'Corporate social responsibility: An economic and financial framework', The Geneva papers on risk and insurance-Issues and practice. Springer, 30(3), pp. 387–409.

Heymans, A., and Brewer, W. P. (2015). The influence of volatility spill-overs and market beta on portfolio construction. South African Journal of Economic and Management Sciences, 18(2), 277–290.

Hill, C. W. L. (1995) 'National institutional structures, transaction cost economizing and competitive advantage: The case of Japan', Organization Science. INFORMS, 6(1), pp. 119–131.

Hoepner, A. G. F. (2010) 'Portfolio diversification and environmental, social or governance criteria: Must responsible investments really be poorly diversified', Available at SSRN 1599334.

Hong, H. and Kacperczyk, M. (2009) 'The price of sin: The effects of social norms on markets', Journal of Financial Economics. Elsevier, 93(1), pp. 15–36.

Hunter, J. E., and Schmidt, F. L. (2004). Methods of meta-analysis: correcting error and bias in research findings (2. ed.). Thousand Oaks, Calif. London.

Ielasi, F., Rossolini, M., and Limberti, S. (2018). Sustainability-themed mutual funds: an empirical examination of risk and performance. The Journal of Risk Finance, 19(3), 247–261.

Ioannou, I., and Serafeim, G. (2012). What drives corporate social performance? The role of nation-level institutions. Journal of International Business Studies, 43(9), 834–864.

Jensen, M. C. (1968). THE PERFORMANCE OF MUTUAL FUNDS IN THE PERIOD 1945–1964. Journal of Finance, 23(2), 389–416.

Kempf, A. and Osthoff, P. (2007). The effect of socially responsible investing on portfolio performance. European Financial Management, 13(5):908–922.

Kennon, J. (2019). How One of the World's Largest Companies Made $3.8 Billion Disappear. Retrieved 1 April 2020, from https://www.thebalance.com/worldcom-s-magic-trick-356121

King, A. and Lenox, M. (2002) 'Exploring the locus of profitable pollution reduction', Management Science. INFORMS, 48(2), pp. 289–299.

Leite, P., and Cortez, M. C. (2018). The performance of European SRI funds investing in bonds and their comparison to conventional funds. Investment Analysts Journal, 47(1), 65–79.

Lintner, J. (1965) 'The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets', Review of Economics and Statistics, Vol. 47, pp. 13–37.

Lipson, M. L., Mortal, S., and Schill, M. J. (2011). On the Scope and Drivers of the Asset Growth Effect. Journal of Financial and Quantitative Analysis, 46(6), 1651–1682.

MacKinlay, A. C. (1995) "Multifactor Models Do Not Explain Deviations from the CAPM," Journal of Financial Economics, 38, 3-28.

Markowitz, H. (1952). PORTFOLIO SELECTION*. Journal of Finance, 7(1), 77–91.

Markowitz, H. (1959) Portfolio selection: Efficient diversification of investments. John Wiley New York.

McWilliams, A. and Siegel, D. (2000) 'Corporate social responsibility and financial performance: correlation or misspecification?', Strategic management journal. Wiley Online Library, 21(5), pp. 603–609.

McWilliams, A. and Siegel, D. (2001) 'Corporate social responsibility: A theory of the firm perspective', Academy of management review. Academy of Management Briarcliff Manor, NY 10510, 26(1), pp. 117–127.

Moskowitz, M. R. (1972). Choosing socially responsible stocks. Business and Society Review, Vol. 1, 71-75.

Mossin, J. (1966) 'Equilibrium in a capital asset market', Econometrica, Vol. 34, 1966, pp. 768–783.

MSCI. (2020). ESG Ratings. Retrieved 22 March 2020, from https://www.msci.com/esg-ratings

National Academy of Sciences (NAS). (2019), National Academy of Engineering (US) and Institute of Medicine (US) Committee on Ensuring the Utility and Integrity of Research Data in a Digital Age. Retrieved 2 March 2020, from: https://www.ncbi.nlm.nih.gov/books/NBK215260/

Oberndorfer, U. et al. (2013) 'Does the stock market value the inclusion in a sustainability stock index? An event study analysis for German firms', Journal of Environmental Economics and Management. Elsevier, 66(3), pp. 497–509.

Ortas, E., Álvarez, I., and Garayar, A. (2015). The Environmental, Social, Governance, and Financial Performance Effects on Companies that Adopt the United Nations Global Compact. Sustainability, 7(2), 1932–1956.

Porter, M. E. (1995) 'Van der Linde (1995)," Green and competitive: ending the stalemate', Harvard Business Review, 73(5), pp. 120–134.

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

Revelli, C., and Viviani, J. (2015). Financial performance of socially responsible investing (SRI): what have we learned? A meta-analysis. Business Ethics: A European Review, 24(2), 158–185.

Robson, C. (2002) Real World Research (2nd edn). Oxford: Blackwell.

Russo, M. V and Fouts, P. A. (1997) 'A resource-based perspective on corporate environmental performance and profitability', Academy of management Journal. Academy of Management Briarcliff Manor, NY 10510, 40(3), pp. 534–559.

Saltaji, I.M.F. (2013). 'Corporate Governance Relation with Corporate Sustainability', Internal Auditing and Risk Management, vol.8, no.2, pp.137-147

Schmidt, P. S. et al. (2015) 'On the construction of common size, value and momentum factors in international stock markets: A guide with applications', CCRS Working Paper Series. University of Zurich, (01/11).

Schmidt, Peter and von Arx, Urs & Schrimpf, Andreas & Wagner, Alexander & Ziegler, Andreas. (2011). On the Construction of Common Size, Value and Momentum Factors in International Stock Markets: A Guide with Applications. SSRN Electronic Journal. 10.2139/ssrn.1738315.

Schramade, W. (2017). Investing in the UN Sustainable Development Goals: Opportunities for Companies and Investors. Journal Of Applied Corporate Finance, 29(2), 87-99. doi: 10.1111/jacf.12236

Schröder, M. (2007) 'Is there a difference? The performance characteristics of SRI equity indices', Journal of Business Finance & Accounting. Wiley Online Library, 34(1-2), pp. 331–348.

Schwert, G. W. (2002). Anomalies and Market Efficiency. Cambridge, Mass., USA.

Shapiro, S. S. and Wilk, M. B. (1965) 'An analysis of variance test for normality (complete samples)', *Biometrika*. JSTOR, 52(3/4), pp. 591–611.

Sharpe, W.F. (1964) 'Capital asset prices: A theory of market equilibrium under conditions of risk', Journal of Finance, Vol. 19, 1964, pp. 425–442.

Sharfman, M. P., and Fernando, C. S. (2008). Environmental risk management and the cost of capital. Strategic Management Journal, 29(6), 569–592.

Statman, M. (2000) 'Socially responsible mutual funds (corrected)', Financial Analysts Journal. Taylor & Francis, 56(3), pp. 30–39.

Stevens, P. (2019). Your complete guide to investing with a conscience, a $30 trillion market just getting started. Retrieved 6 March 2020, from https://www.cnbc.com/2019/12/14/your-complete-guide-to-socially-responsible-investing.htm

Thomson Reuters. (2015). Thomson Reuters Global Equity Indices - Index Methodology. Retrieved 22 March 2020, from https://www.thomsonreuters.com/content/dam/openweb/documents/pdf/tr-com-financial/methodology/global-equity-index-methodology-oct-2015.pdf

Thomson Reuters. (2019). Environmental, Social and Governance (ESG) Scores from Refinitiv. Retrieved 22 March 2020, from https://www.refinitiv.com/content/dam/marketing/en_us/documents/methodology/esg-scores-methodology.pdf

Tirole, J. (2001) 'Corporate Governance', Econometrica, 69(1), pp. 1–35.

Treynor, J.L. (1962) ‘Toward a theory of market value of risky assets’, in R.A. Korajczyk (ed.), Asset pricing and portfolio performance. London: Risk Publications.

Ullmann, A. A. (1985). Data in Search of a Theory: A Critical Examination of the Relationships among Social Performance, Social Disclosure, and Economic Performance of U. S. Firms. The Academy of Management Review, 10(3), 540–557.

United Nations. (2014). ESG screening in fixed income investing. Retrieved 6 April 2020, from https://www.unpri.org/fixed-income/esg-screening-in-fixed-income-investing/36.article

United Nations. (2020). Sustainable Development Goals .:. Sustainable Development Knowledge Platform. Retrieved 7 April 2020, from https://sustainabledevelopment.un.org/?menu=1300

United Nations. (2020). What are the Principles for Responsible Investment?. Retrieved 6 April 2020, from https://www.unpri.org/pri/an-introduction-to-responsible-investment/what-are-the-principles-for-responsible-investment

Van de Velde, E., Vermeir, W., and Corten, F. (2005). Corporate social responsibility and financial performance. Corporate Governance: The International Journal of Business in Society, 5(3), 129–138.

Waddock, S. A. and Graves, S. B. (1997) ‘The corporate social performance–financial performance link’, Strategic management journal. Wiley Online Library, 18(4), pp. 303–319.

Walley, N., and Whitehead, B. (1994). It’s Not Easy Being Green. Harvard Business Review. 72, 2-7 Retrieved 10 April 2020, from https://hbr.org/1994/05/its-not-easy-being-green

White, M. (1996). Corporate Environmental Performance and Shareholder Value. McIntire School of Commerce Retrieved 13 March 2020, from https://libraopen.lib.virginia.edu/downloads/707957685

Wooldridge, J. M. (2016) *Introductory econometrics: A modern approach*. Nelson Education.

Yamashita, M., Sen, S., and Roberts, M.C. (1999). THE REWARDS FOR ENVIRONMENTAL CONSCIENTIOUSNESS IN THE U.S. CAPITAL MARKETS.

Zhang, L. (2017). The Financial Return of Responsible Investing. Retrieved 11 February 2020, from https://spilplatform.com/wp-content/uploads/sites/329/2017/02/SPIL-The-Financial-Return-of-Responsible-Investing.pdf

Zhang, L. (2017). The Investment CAPM. European Financial Management, 23(4), 545–603.

Ziegler, A., Busch, T. and Hoffmann, V. H. (2011) ‘Disclosed corporate responses to climate change and stock performance: An international empirical analysis’, Energy Economics. Elsevier, 33(6), pp. 1283–1294.

11 Table of Illustrations

List of Equations:

List of Figures:

List of Tables:

12 Appendix

12.1 Influential Literature Analysis

Research area characteristics	
1. Research area maturity	Is the research area mature enough?
2. Research area size	Is the sufficient literature in the research area to warrant a comprehensive literature survey?
Research synthesis preparation	
3. Research area definition coherenece	Is the literature surveyed coherently bounded?
4. Study assessment process	Does the review provide a good indication of the key papers in the research area?
5. Study selection process	Are there justifiable reasons why certain literature is included while other literature is excluded?
Research synthesis conduct	
6. Quality of analysis with respect to historical evolution	Does the article provide a thorough insight into where the literature is now, and why?
7. Quality of analysis with resprect to methodologies	Is the analysis of the literature suveyed complete, in terms of discussions of any contrasting methodologies used in the literature?
8. Quality of analysis with respect to findings and views	Is the analysis of the literature surveyed complete in terms of discussing agreements, disagreements, and contrasting views?
9. Quality of conclusions	Does it draw reasoned and authoritative conclusions as to where the literature is/ should be heading?
10. Quality of suggestions for further research	Does the review draw reasoned and authoritative conclusions? What are the important questions left to be asked?
Research synthesis customisation	
11. Clarity of review	Will the literature review be sensibly understood by its intended audience?

Explanation: Table 1 displays the assessment criteria for literature reviews indicated by Armstrong and Wilkinson (2005;2007). The second column shows the questions stated by Armstrong and Wilkinson as assessment criteria for referees of the International Jounal of Management Reviews, the dominant reviews journal in the wider area of Management currently edited by them, while the first column provides the assessment criteria title assigned by Hoepner and McWilliams to the respective question. The four terms (see blue shading) represents group titles for the assessment criteria by Hoepner and McWilliams (2009).

12.2 Influential Literature Overview

The next section is devoted to give an overview of the most influential literature that have been used for the study.

Title	Author(s)	Publisher	Year	Category	Data and Models	Conclusions
The wages of social responsibility- where are they? A critical view of ESG investing	(Halbritter & Dorfleitner)	Review of Financial Economics	2015	Stocks	**Source(s):** Asset4, KLD, Bloomberg **Model(s):** CAPM, three factor model. **Scope:** The relationship between ESG-scores and stock returns (Sub-periods).	ESG portfolios with low and higher ESG scores show no significant difference in returns, for both the total ESG score and the individual parameters. The study finds that the magnitude is substantially dependent on the individual rating provider used, the company sample and the particular sub-period. The result suggest that investors should no longer expect abnormal returns with regards to trading portfolios with high and low rated firms with regard to ESG aspects.
Corporate Social Responsibility and Credit Ratings	(Attig, Ghoul, Guedhami, & Suh)	Journal of Business Ethics	2013	Credit Ratings	**Source(s):** Thomson IBES, Compustat, KLD **Model(s):** Regression, OLS, Robustness test **Scope:** Firms' financing costs and shareholder value (S&P 500).	The study finds a significant positive relationship between CSR and corporate credit ratings. Furthermore, the rating agencies tends to award relatively high ratings to firms with good social performance. Investing in CSR initiatives leads to lower financing costs, enhance firm value and increased shareholder value. Significant positive impact on both aggregate CSR score and individual CSR score.
Does the stock market value the intrinsic inclusion in a sustainability stock index? An event study analysis for German firms	(Oberndorfer, Schmidt, Wagner, & Ziegler)	Journal of Environmental Economics and Management	2013	Stock indices	**Source(s):** DJSI Stoxx, DJSI World, Dow Jones, STOXX limited, SAM Group **Model(s):** CAPM, three factor model, t-GARCH **Scope:** Effect on stock performance by incl. in ESG stock indexes.	The empirical results suggest that the stock market punish firms included in ESG screened indices. The study notes that the criteria for inclusion varies between indices, which reduces the comparability. The findings include a strongly negative effect in the DJSI World, and no significant average cumulative abnormal return for companies in DJSI Stoxx. This implies that the inclusion in a more visible and used sustainability stock index may have larger impacts.
Disclosed corporate responses to climate change and stock performance	(Ziegler, Busch, & Hoffmann)	Energy Economics	2011	Stocks	**Source(s):** Asset4 **Model(s):** CAPM, four factor model **Scope:** Environmental initiatives and stock performance in Europe and the US.	The study compares portfolios with varying degrees of environmental performance. In general, they find an insignificant relationship between firms' environmental initiative and their stock returns. However, they find that the relationship between disclosed corporate responses to climate change and stock performance have been positive for energy firms in the US.

Title	Author(s)	Publisher	Year	Category	Data and Models	Conclusions
The eco-efficiency premium puzzle	(Derwall, Guenster, Bauer, Koedijk)	*Financial Analysts Journal*	2005	Stocks	**Source(s):** Innovest **Model(s):** CAPM, three factor model, four factor model **Scope:** Different portfolios with high/low ranked companies	The study finds that portfolios made with companies with high environmental scores performs better that companies with low environmental scores. Difference on the portfolios were also evident for risk-adjusted return. The portfolio with companies with a low environmental score has a significantly lower Sharpe ratio than the market proxy.
Portfolio diversification and environmental, social or governance criteria: Must responsible investments really be poorly diversified?	(Hoepner)	*Principles for Responsible Investment, PRI Secretariat*	2010	Stocks	**Source(s):** Results from previous studies **Model(s):** Statman diversification **Scope:** Diversification of ESG-portfolios	The study concludes that portfolios which are diversified using stock with low ESG scores, in an attempt to minimize risk, in reality results in the opposite due to increased ESG risk associated with the stock. The study finds that the portfolios with highest average ESG score has lower total risk, when two portfolios which have identical exposure to the diversification drivers "number of stocks" and "correlation between selected stocks" are compared to each other.
Socially responsible investments: institutional aspects, performance, and investor behaviour	(Renneboog, Horst, & Zhang)	*Journal of Banking & Finance*	2008	Stocks	**Source(s):** KLD **Model(s):** CAPM, four factor **Scope:** Study SRI and whether the market is in equilibrium or if there is an out-performance of SR stocks	The study finds that although CSR are found to be associated with shareholder value, there is no convincing evidence on the direction of the causality. In addition, they find no evidence that environmental performance or good corporate governance produces superior abnormal returns. They also argue that stock markets misprice information on CSR on the short term.
The stocks at stake: Return and risk in socially responsible investment	(Galema, Plantinga, & Scholtens)	*Journal of Banking & Finance*	2008	Stocks	**Source(s):** KLD, Thomson Reuters **Model(s):** Four factor model,Pooled OLS with robust standard deviation **Scope:** Stock return	The study concludes with that aggregate analysis of SRI scores may eliminate a relationship if individual dimensions of SRI have opposite effects on performance. Furthermore, the study finds that SRI portfolios that score positive on diversity and environment has a significant impact on returns.

Title	Author(s)	Publisher	Year	Category	Data and Models	Conclusions
Environmental risk management and the cost of capital	Sharfman & Fernando)	*Strategic Management Journal*	2008	Stocks	**Source(s):** KLD, EPA TRI Data from IRRC, Bloomberg **Model(s):** Fama & French, CAPM, ANOVA, Colinearity diagnostics and hierarchical regression analysis **Scope:** Portfolios in the US (S&P 500)	The study finds that improved environmental risk management is associated with a lower cost of capital, particularly through the lowering of the volatility of the firm's stock as measured by beta. The results suggest that environmental risk management are rewarded by the equity markets and thus give a lower cost of equity due to the fact that more individuals purchase the high environmental performer's stock. In addition, the cost of foreign capital seemed to be higher for socially responsible companies, meaning that these companies attract attention internationally.
Mutual Fund Attributes and Investor Behavior	(Bollen)	*Journal of Financial and Quantitative Analysis*	2007	Funds	**Source(s):** CRSP database (mutual fund) **Model(s):** CAPM, four factor model **Scope:** Funds' cash flow volatility and the investor's utility function	The study concludes with that the monthly volatility of investor cash flow is lower in socially responsible funds than in conventional funds. In addition, there is evidence that cash outflows from socially responsible funds are less sensitive to lagged negative returns. These results imply that investors derive utility from socially responsible attributes, when returns are positive.
International evidence on ethical mutual fund performance and investment style	(Bauer, Koedijk, & Otten)	*Journal of Banking & Finance*	2004	Funds	**Source(s):** Morningstar (US), EIRIS (UK), Ecoreporter, Datastream **Model(s):** CAPM, four factor model **Scope:** Funds and indices	The study finds no evidence of a statistically significant difference in return between ethical and conventional mutual fund. This is after controlling for size, book-to-market and momentum. Next, ethically funds are typically less exposed to market return variability compared to conventional funds. Ethical funds seem to be more growth-oriented, than value-oriented.
Sustainable investing: A "why not" moment. Environmental, social and governance investing insights	(Blackrock)	*Global Insights*	2018		**Source(s):** Asset4, Thomson Reuters **Model(s):** Low volatility, quality, value, momentum **Scope:** ESG & Contr.	The study finds that companies with most ESG policies are more involved in controversies, while companies with few ESG policies had few controversies. Size (large companies are involved in controversy the most), existing controversies and number of policies are the most important variables that contributes to future controversies.

Title	Author(s)	Publisher	Year	Category	Data and Models	Conclusions
ESG and financial performance: aggregated evidence from more than 2000 empirical studies	(Friede, Busch, & Bassen)	*Journal of Sustainable Finance & Banking*	2015	Meta-analysis: Funds and Portfolios	**Source(s):** Multiple **Model(s):** Multiple, including substantial amount of factor models **Scope:** ESG and financial performance of funds and portfolios across in North America, Europe and emerging markets	As the knowledge on the relationship between financial return and ESG performance is fragmented, the study aims to overcome this shortcoming by gathering a significant amount of research. The study combines 2.200 individual studies and finds that a substantial amount of studies finds a non-negative relation. However, the results are highly dependent on geography, with North America having the most positive relationships. Only a minority (26,1%) of European studies find a positive relationship. More importantly, their findings appear to be stable over time.
The Financial Return of Responsible Investing	(Zhang)	*Sustainable Pension Investment Lab*	2017	Literature study: Funds and portfolios	**Source(s):** Multiple **Model(s):** Multiple, including substantial amount of factor models **Scope:** Responsible portfolio performance vs conventional benchmark	Overall conclusion is that responsible investments do not sacrifice financial returns. On the other hand, it may help reduce risk exposure. ESG could be considered a risk factor, and if it is priced in, ESG high performers may produce negative risk-adjusted excess returns, whereas the low performers may experience the opposite.
Financial performance of socially responsible investing (SRI): what have we learned? A meta-analysis.	(Revelli, & Viviani)	*Business Ethics: A European Review*	2015	Meta-analysis:	**Source(s):** Multiple **Model(s):** Multiple, including substantial amount of factor models **Scope:** Responsible portfolios vs. conventional portfolios	The results indicate that CSR considerations in portfolios is neither an advantage or disadvantage with respect to financial performance. Results vary as methodology and applied market benchmark also differ between studies. Hence, the authors conclude that results are highly dependent on these underlying factors.
Do ESG controversies matter for firm value? Evidence from international data	(Aouadi, & Marsat)	*Journal of Business Ethics*	2018	Portfolios	**Source(s):** Asset4 **Model(s):** Panel data analysis, Tobin's Q as dependent variable. **Scope:** Relationship between ESG controversies and company value.	The authors examine more than 4.000 companies from 58 countries from 2002-2011. Their primary analysis show that ESG controversies are associated with relatively higher market value of firms. Further, they find that Corporate Social Performance (CSP) only has an effect on market value for high-attention companies, as these are already highly developed and operated in developed economies with greater transparency.

12.3 Residual Plots, ESG Portfolios

ESG High: F&F5 + ESG Factor

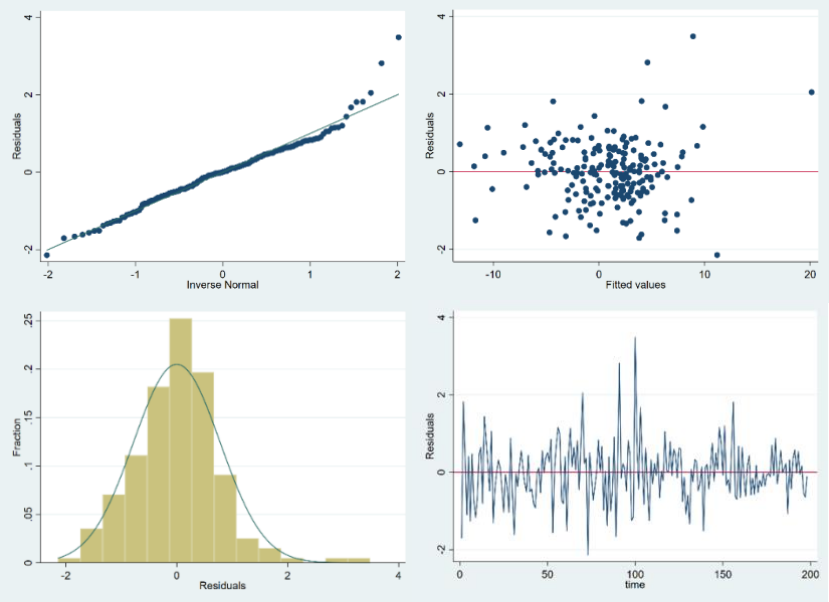


Figure 12: The figure displays four variations of residual plots for the ESG High portfolio against F&F5 and the self-constructed ESG Factor. The figures are Q-Q Normality plot, residuals against fitted values, residual histogram and residual over time. The regression is displayed in model 5 in Table 7.

ESG High: F&F5 + Controversy Factor

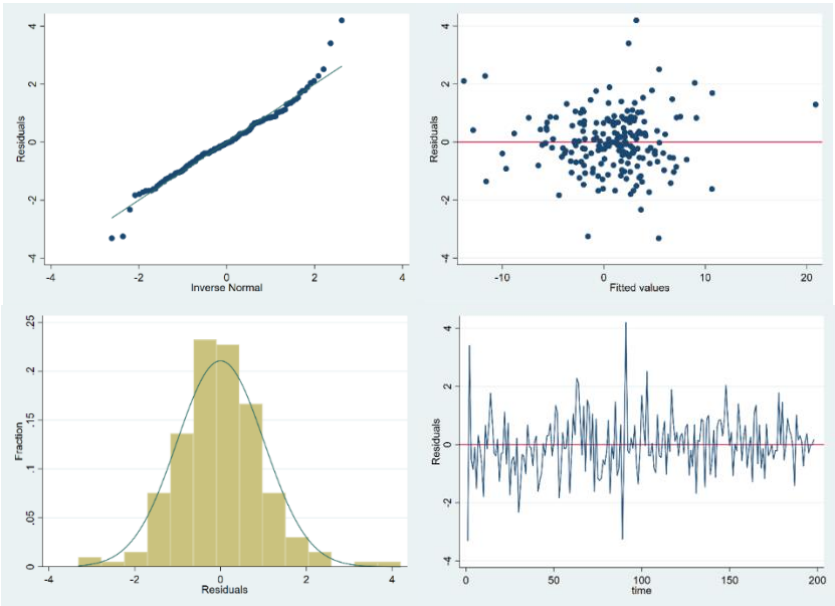


Figure 13: The figure displays four variations of residual plots for the ESG High portfolio against F&F5 and the self-constructed Controversy Factor. The figures are Q-Q Normality plot, residuals against fitted values, residual histogram and residual over time. The regression is displayed in model 6 in Table 7.

ESG Low: F&F5 + ESG Factor

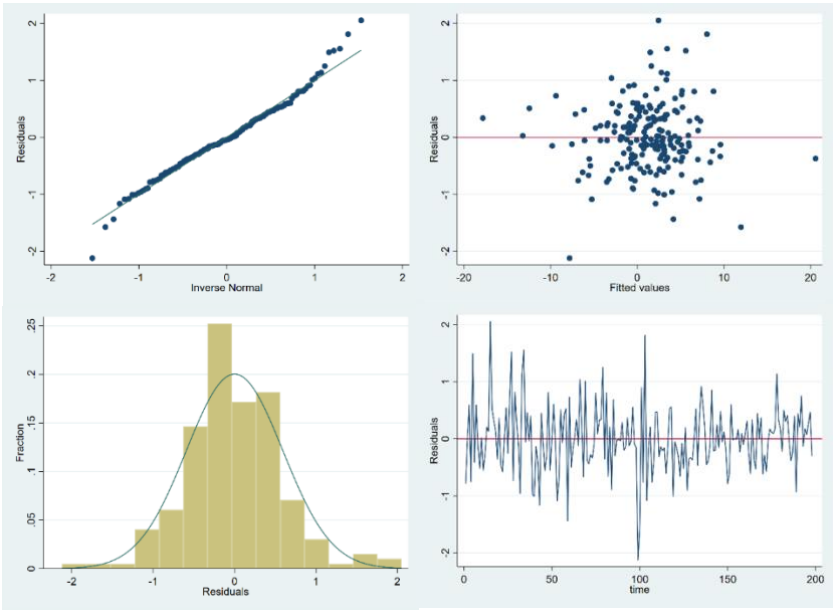


Figure 14: The figure displays four variations of residual plots for the ESG Low portfolio against F&F5 and the self-constructed ESG Factor. The figures are Q-Q Normality plot, residuals against fitted values, residual histogram and residual over time. The regression is displayed in model 5 in Table 7.

ESG Low: F&F5 + Controversy Factor

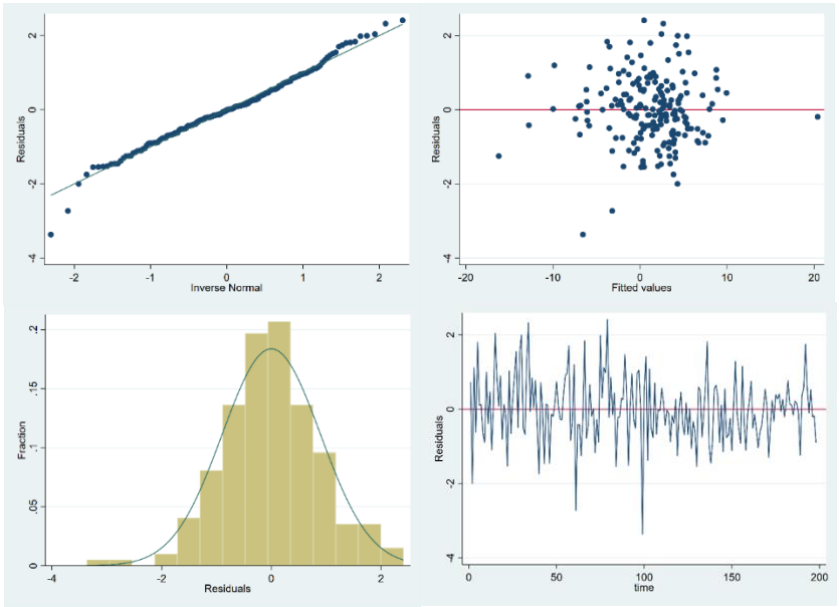


Figure 15: The figure displays four variations of residual plots for the ESG Low portfolio against F&F5 and the self-constructed Controversy Factor. The figures are Q-Q Normality plot, residuals against fitted values, residual histogram and residual over time. The regression is displayed in model 6 in Table 7.

ESG Residuals

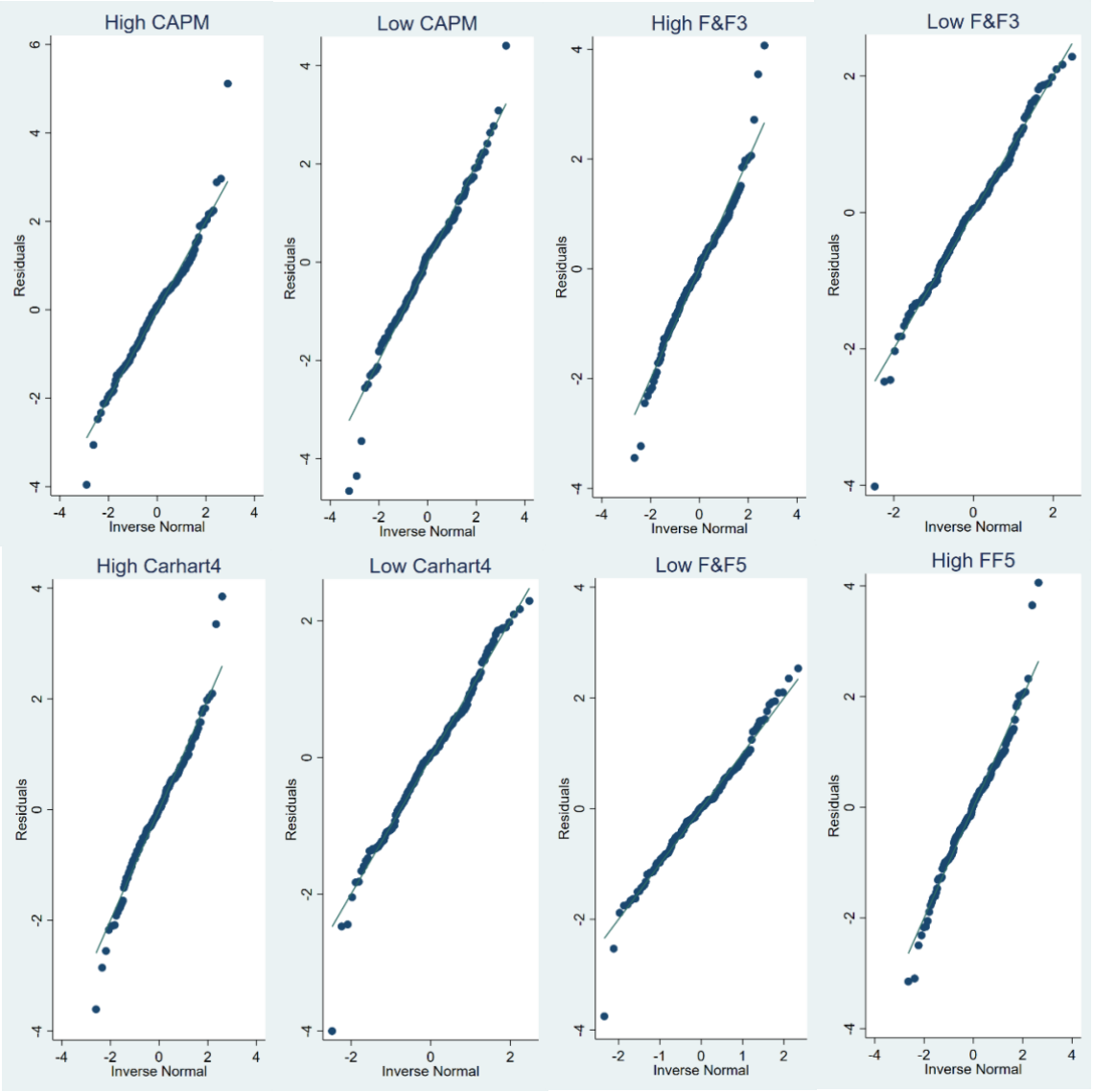


Figure 16: ESG Residuals for regressions in Table 7.

12.4 Residual Plots, Controversy Portfolios

Controversy High: F&F5 + ESG Factor

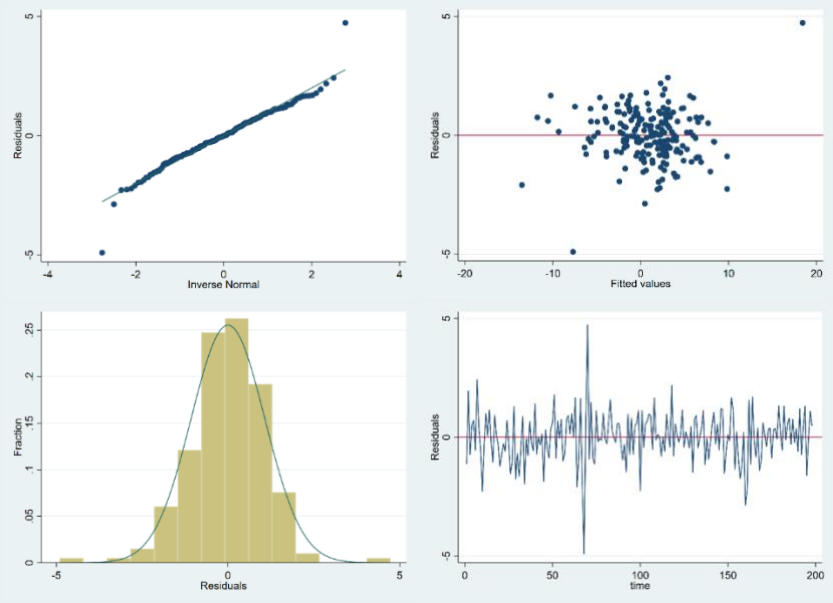


Figure 17: The figure displays four variations of residual plots for the Controversy High portfolio against F&F5 and the self-constructed ESG Factor. The figures are Q-Q Normality plot, residuals against fitted values, residual histogram and residual over time. The regression is displayed in model 5 in Table 8.

Controversy High: F&F5 + Controversy Factor

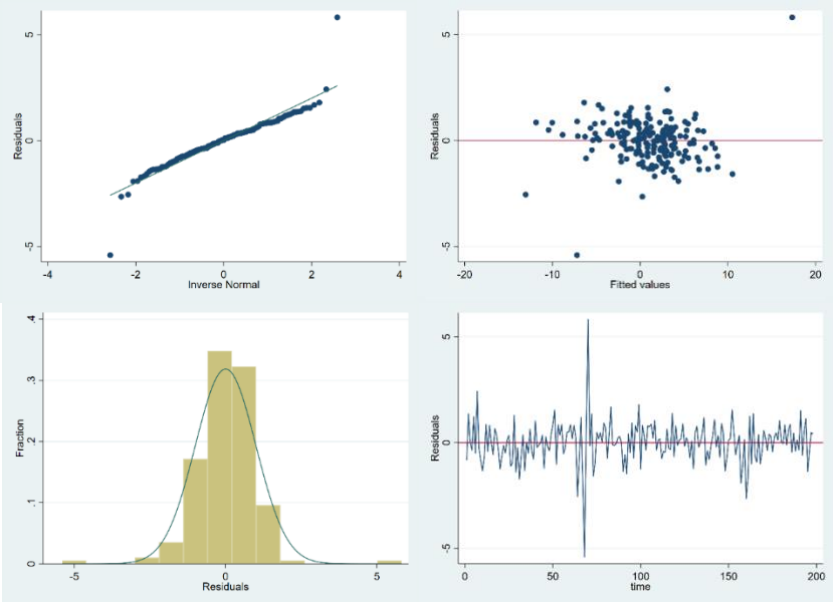


Figure 18: The figure displays four variations of residual plots for the Controversy High portfolio against F&F5 and the self-constructed Controversy Factor. The figures are Q-Q Normality plot, residuals against fitted values, residual histogram and residual over time. The regression is displayed in model 6 in Table 8.

Controversy Low: F&F5 + ESG Factor

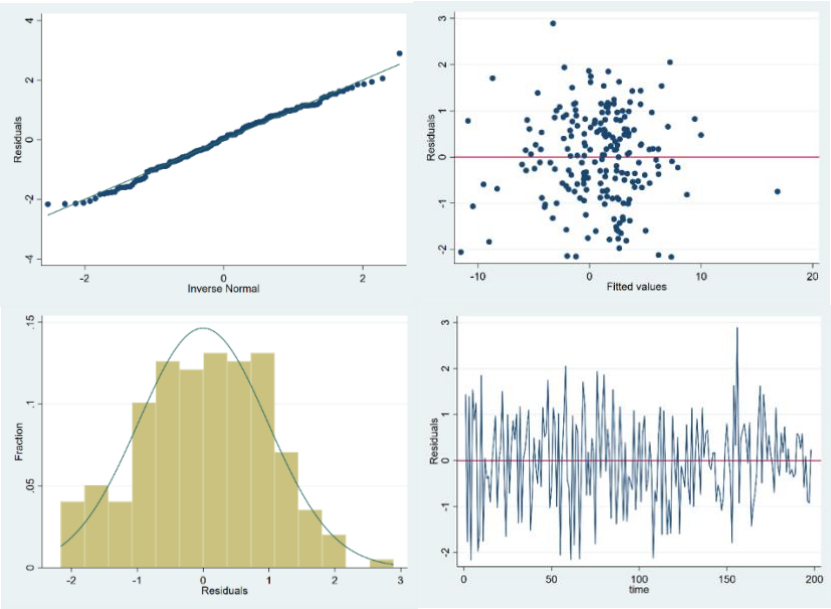


Figure 19: The figure displays four variations of residual plots for the Controversy Low portfolio against F&F5 and the self-constructed ESG Factor. The figures are Q-Q Normality plot, residuals against fitted values, residual histogram and residual over time. The regression is displayed in model 5 in Table 8.

Controversy Low: F&F5 + Controversy Factor

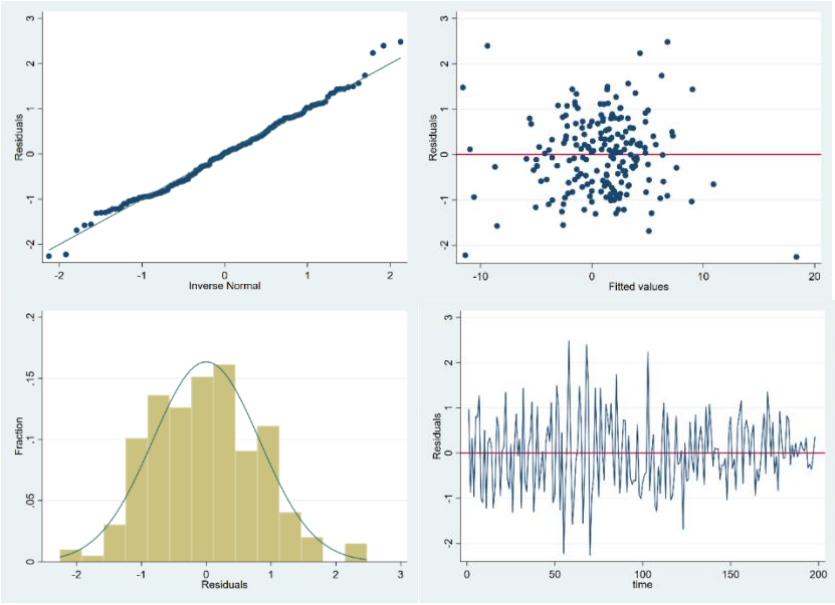


Figure 20: The figure displays four variations of residual plots for the Controversy Low portfolio against F&F5 and the self-constructed Controversy Factor. The figures are Q-Q Normality plot, residuals against fitted values, residual histogram and residual over time. The regression is displayed in model 6 in Table 8.

Controversy Residuals

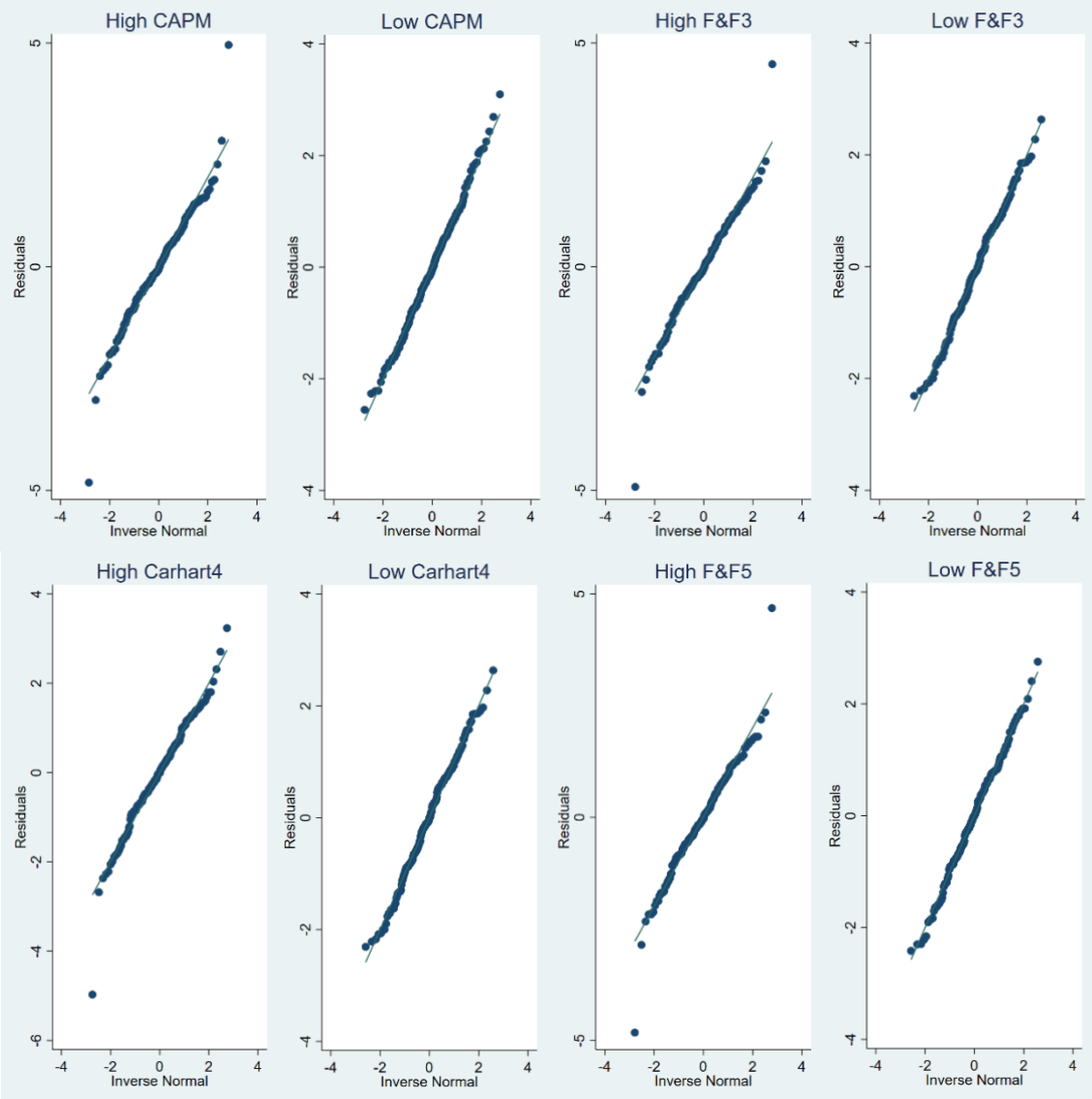


Figure 21: Controversy residuals for regressions in Table 8.

12.5 Linearity Plots

Linearity- ESG High

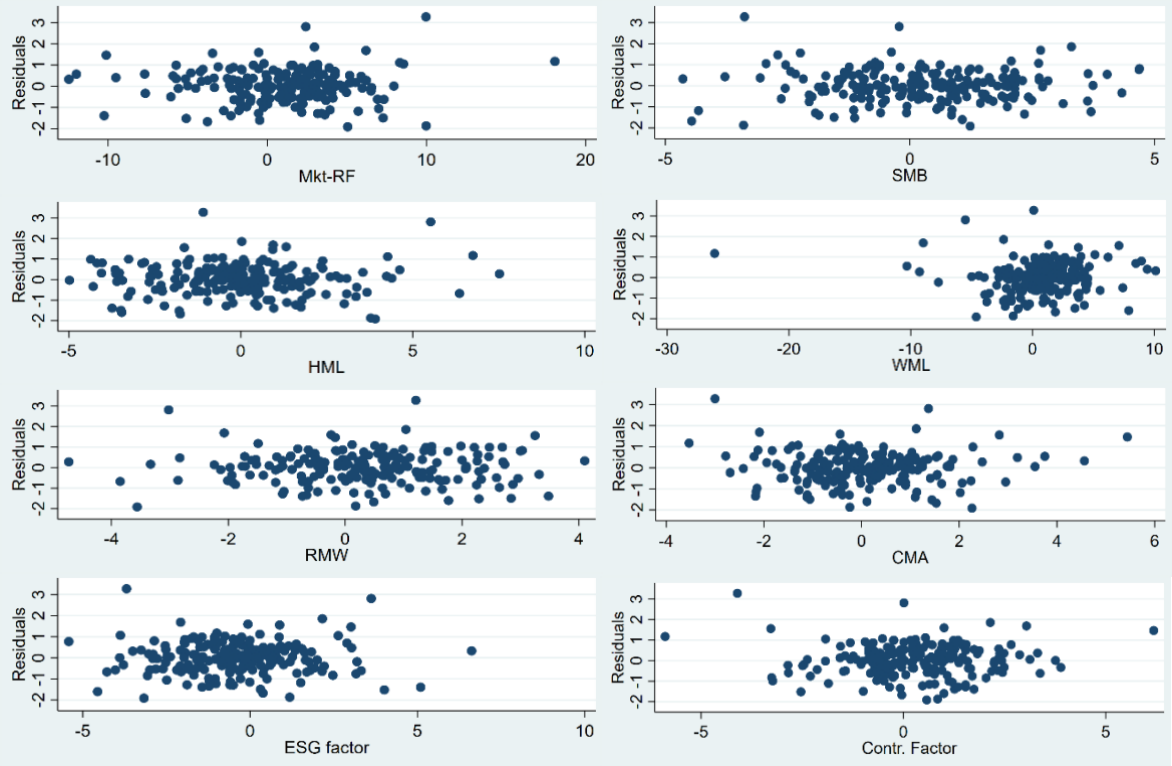


Figure 22: The figure depicts linearity plots. The residuals are plotted against the various independent variables.

Linearity- ESG Low

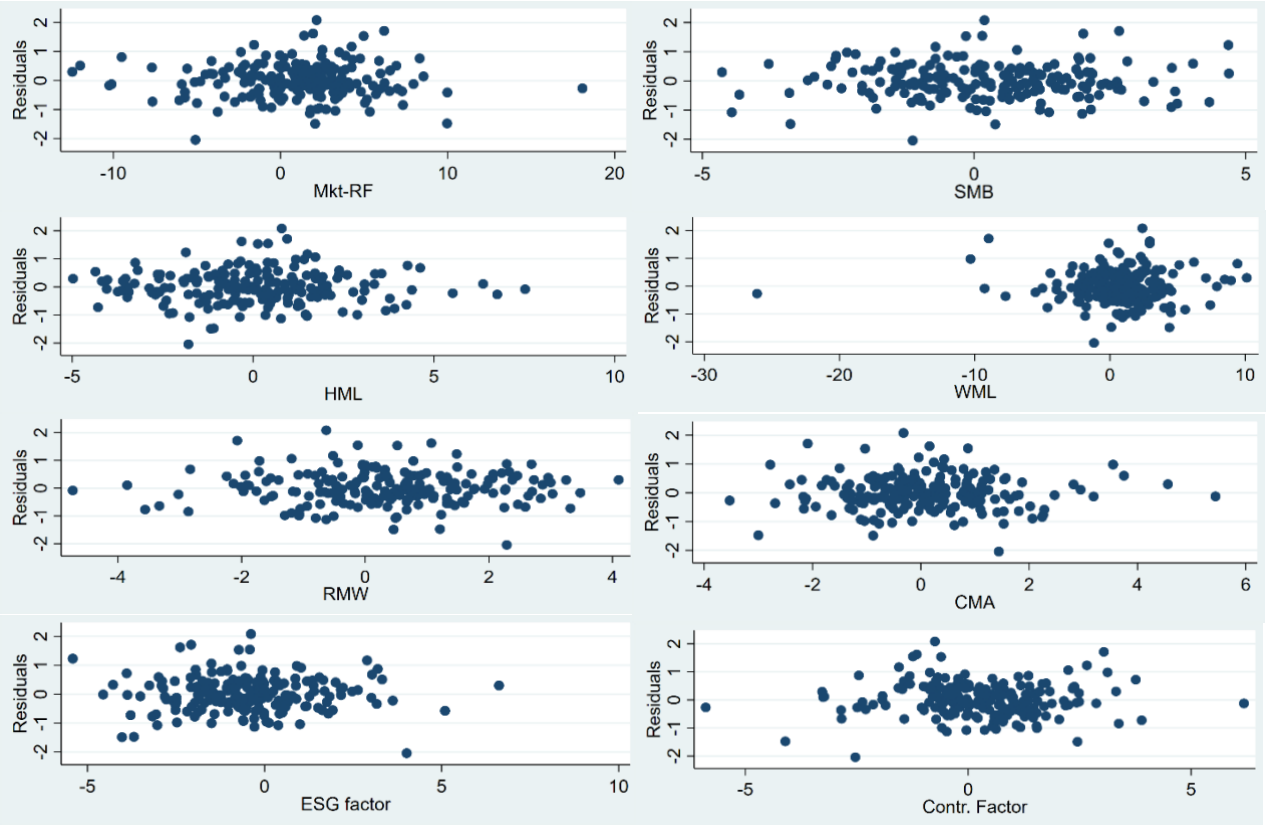


Figure 23: The figure depicts linearity plots. The residuals are plotted against the various independent variables.

Linearity- Controversy High

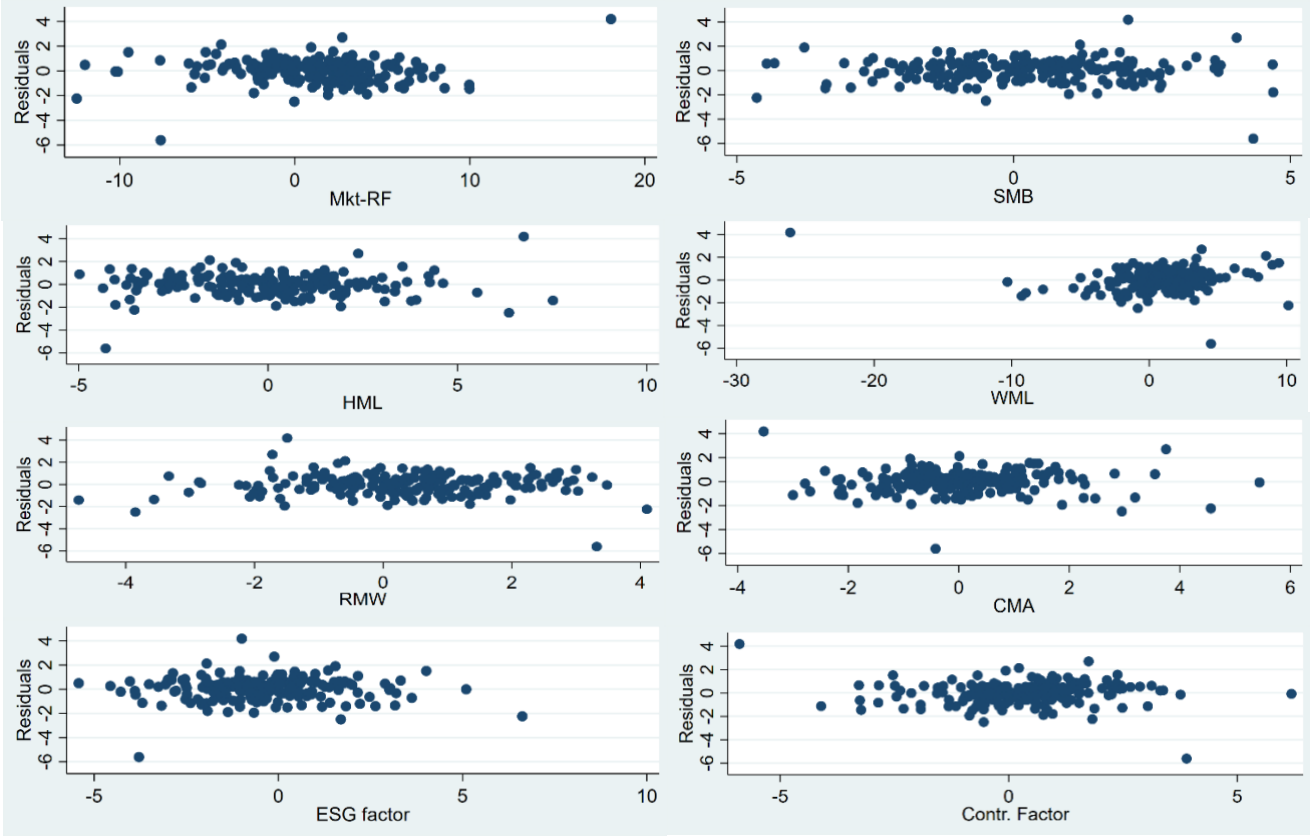


Figure 24: The figure depicts linearity plots. The residuals are plotted against the various independent variables.

Linearity- Controversy Low

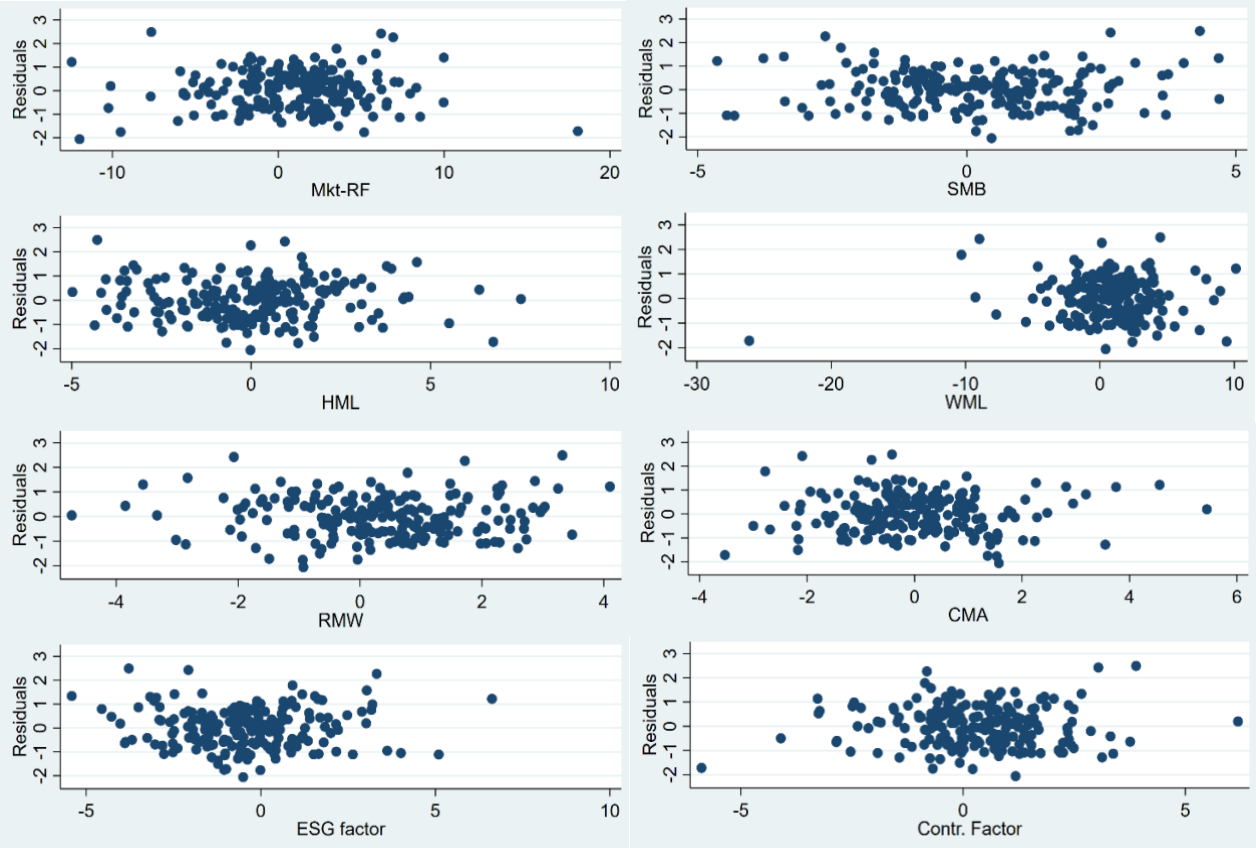


Figure 25: The figure depicts linearity plots. The residuals are plotted against the various independent variables.

12.6 Model testing

Breusch-Pagan/ Cook-Weisberg test for heteroskedasticity

Test for heteroskedasticity in ESG portfolios

	ESG Portfolios				Controversy Portfolios			
	ESG Low		ESG High		Controversy Low		Controversy High	
Models	chi^2	P-value	chi^2	P-value	chi^2	P-value	chi^2	P-value
CAPM	17,600	0,000	6,470	0,011	0,070	0,798	3,730	0,053
Three-factor	35,890	0,000	20,730	0,000	7,740	0,052	9,580	0,023
Carhart	7,990	0,092	9,970	0,410	7,740	0,1015	11,17	0,025
Five-Factor	5,640	0,342	19,010	0,002	13,310	0,021	2,800	0,094
Five-Factor + ESG	4,150	0,656	33,240	0,000	10,020	0,124	25,680	0,003
Five Factor + Controversy	7,710	0,260	15,420	0,017	3,740	0,712	75,810	0,000

Table 12: The table depicts the results from the Breusch-Pagan test for heteroskedasticity, for both ESG and Controversy portfolios. H_0 for the test is presence of homoskedasticity. Hence, a P-value above the 5% threshold implies that the model does not have problems with heteroskedasticity. As can be read from the table, several of the regressions have heteroskedastic problems. Accordingly, to handle this issue, robust standard errors will be applied.

Breusch-Godfrey test for autocorrelation

Test for autocorrelation in portfolios

	ESG Portfolios				Controversy Portfolios			
	ESG Low		ESG High		Controversy Low		Controversy High	
Models	chi^2	P-value	chi^2	P-value	chi^2	P-value	chi^2	P-value
CAPM	16,537	0,168	11,520	0,485	14,691	0,259	18,938	0,090
Three-factor	7,173	0,846	8,825	0,718	13,929	0,305	16,140	0,185
Carhart	7,116	0,850	15,347	0,223	13,959	0,303	15,817	0,200
Five-Factor	7,514	0,822	9,060	0,698	14,614	0,263	16,648	0,163
Five-Factor + ESG	6,168	0,907	11,343	0,500	13,216	0,354	17,618	0,128
Five Factor + Controversy	5,645	0,933	10,473	0,577	18,983	0,089	9,630	0,648

Table 13: The table depicts the results from the Breusch-Godfrey test for autocorrelation, for both ESG and Controversy portfolios. H_0 for the test is that autocorrelation is not present in the regression model. Hence, if a high Chi^2-value and a P-value below the 5% threshold is obtained, the model has issues with autocorrelation. In this case, the P-values is above the statistical threshold. Hence, the authors fail to reject the null hypothesis and concludes that there may not be autocorrelation present.

Augmented Dickey Fuller test for stationarity

Test for stationarity in portfolios		
	Test values	
Parameters (independent)	Z (t)	P-value
ESG High, Excess return	-10,557	0,000
ESG Low, Excess return	-8,813	0,000
Controversy High, Excess return		
Controversy Low, Excess return		
Parameters (dependent)	Z (t)	P-value
Mrkt-Rf	-9,611	0,000
SMB	-9,874	0,000
HML	-10,000	0,000
RMW	-9,829	0,000
CMA	-7,396	0,000
ESG Factor	-11,013	0,000
Controversy Factor	-10,011	0,000

Table 14: The table depicts the results from the Augmented Dickey-Fuller test for stationarity for each respective dependent and independent variable applied in the analysis. The test's H_0 is that a unit root is present, i.e. non-stationarity. Hence, a low absolute |Z(t)| implies a stationarity issue. According to the clear rejection of H_0, the authors conclude that the variables are in fact stationary and thus suitable for regressions.

VIF test for multicollinearity

VIF test		
Variable	VIF	1/VIF
HML	3,630	0,276
RMW	2,980	0,335
ESG Score	1,700	0,590
Mrkt-Rf	1,660	0,276
CMA	1,660	0,601
SMB	1,490	0,672
WML	1,400	0,716
Controversy Score	1,210	0,823
Mean VIF	1,966	

Table 15: The variance inflation factor of a variable is identical to the ratio of overall model variance to the variance of a regression model which only includes the respective independent variable. The ratio has been calculated for each variable applied in the analysis. A high VIF is indicative of collinearity between the respective variable and other variables in the model. As a rule of thumb, a VIF above 10 implies severe collinearity problems. The results of the VIF test show that it is well below this limit.

Shapiro-Wilk test for normality

Shapiro Wilk test				
Variables	W	V	Z	P-value
Mrkt-Rf	0,96074	5,806	4,045	0,00003
HML	0,99546	0,672	-0,914	0,81974
RMW	0,98133	2,762	2,336	0,00975
WML	0,98977	1,512	0,951	0,1707
CMA	0,97284	4,017	3,198	0,00069
ESG Factor	0,98287	2,534	2,138	0,01626
Controversy Factor	0,97413	3,825	3,085	0,00102

Table 16: The table depicts the results from the Shapiro-Wilks test for normality for each respective independent variable applied in the analysis. The test's H_0 is that the data is normally distributed. If the p-value is below the chosen 5% threshold, the H_0 that the data is normally distributed is rejected. If the p-value is greater than the 5% threshold, the H_0 is not rejected. As seen in the table, the results produced for the variables differ. How this is handled is discussed in 5.3.