UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

Graduate School

Are ESG investments a sustainable investing strategy?

Master of Science in Finance

University of Gothenburg, School of Business, Economics and Law
Institution: Center for Finance
Supervisor: Viktor Elliot
Master thesis, Spring 2021
30 credits

Sofia Hallberg and Hanna Johnsson

Gothenburg, Sweden
June 2021

Abstract

Sustainable investments are rapidly growing, and screening of Environmental, Social, and Governance (ESG) scores is a popular method to assess sustainable companies. With the increased usage of this method, the question arises; is screening of ESG scores a sufficient method to allocate investments to companies with top sustainable performance? This study investigates the relation between ESG scores and sustainable performance in the S&P 500 over five years from January 2015 to December 2019. Four ESG portfolios are constructed to examine sustainable performance; two portfolios consisting of companies with high ESG scores and two portfolios with low ESG scores. These portfolios are compared on 10 sustainability metrics. In addition, the relationship between the ESG score and stock return is investigated using the Carhart four-factor model. We find significant support that the top ESG portfolios perform better on most sustainability metrics, although not on all. The results imply that screening of ESG scores is a sufficient method to allocate investments to companies with top sustainable performance. Hence, it can be used for investors to make informed investment decisions and contribute to a more sustainable society. Further, we find a negative relation between ESG score and stock return; the coefficient for the ESG score is significant and negative, although relatively small, and the top portfolios perform slightly worse than the bottom portfolios when comparing the risk-adjusted alphas.

Key words

ESG investing, Sustainable performance, Sustainability, Financial performance

Acknowledgements

We want to express our gratitude to our supervisor Viktor Elliot for his guidance, ideas, and valuable introductions. We thank Zelalem Abay for continuously helping us with various questions related to the data. We also want to express a special thanks to Ylva Baeckstrom and the team at MDOTM, namely Alessio Fiore, Giacomo Barbanti, and Federica Biagi, for their valuable input and discussions throughout the thesis work. Finally, we thank each other for great collaboration, with respect and understanding, throughout the writing process.

Responsibility for any remaining errors lies with the authors alone.

Table of Contents

1. Introduction

Environmental and social issues have gained increased attention and coverage in the media, and consequently increased public awareness, during the last years. In 2015, the member states of the United Nations adopted the 2030 Agenda for Sustainable Development, which contains 17 goals to achieve sustainable development, commonly referred to as the "Sustainable Development Goals (SDGs)" (United Nations, 2021). The financial sector has an important role in the transition into a more sustainable society because it can direct funds toward the most productive investments after taking the social costs and benefits into full consideration (Shiller, 2013). During the last decade, sustainable investments have increased drastically (Boffo & Patalano, 2020), and, following the introduction of the SDGs, ESG (Environmental, Social, and Governance) became a buzzword in the financial industry (PwC, 2020). The percentage of global investors that apply an ESG strategy has increased from 48 percent in 2017 to 75 percent in 2019 (Collins & Sullivan, 2020). Further, ESG mandated assets in the US grow almost three times as fast as non ESG mandated assets, and Collins and Sullivan (2020) predicted that half of all the managed assets in the US could be ESG mandated by 2025. According to Vanwalleghem and Mirkowska (2020), the global increase in sustainable investments can serve as a catalyst to stimulate the transition to a more sustainable society. Hence, a viable method must exist to appropriately target companies with a positive, sustainable impact. One of the most commonly used measurements for corporate sustainability is the ESG score, and investors rely on the ratings when undertaking financial decisions (Berg, Koebel & Rogibon, 2020b). When ESG investing moves into the mainstream arena, the focus shifts from financial performance to sustainable performance (Bos, 2018). With this shift, investors seek to answer questions such as: To what extent does an investment portfolio contribute to a more sustainable society, how much CO2 emission does the portfolio generate, and does the portfolio contribute to the SDGs (Bos, 2018)? These questions call for further research.

The main objective of our research is to investigate whether investments in companies with high ESG scores contribute to a more sustainable society. Specifically, our research question is formulated as follows:

Is screening of ESG scores a sufficient method to allocate investments to companies with top sustainable performance?

The research question is further explored through hypothesis testing, as described in the *Hypotheses section*. It is, however, essential to maintain the financial aspect, as financial performance attracts investments. Hence, if there is no financial incentive to engage in ESG investments, the potential sustainability performance of companies with high ESG scores could be less relevant. We therefore also test if there is a relation between ESG score and financial performance.

With this background, we aim to add an exploration of sustainable performance to the already existing literature on the relation between ESG score and financial performance. More specifically, to extend previous research in the area, we explore sustainable performance and test whether companies with high ESG scores perform better than companies with low ESG scores on a range of sustainability variables, such as CO2 emission, board gender diversity and net employment creation. The sustainability variables are partly chosen based on the UN SDGs to discuss how effective the ESG score is to assess companies working in line with the goals (for further description of the sustainability variables, see the *Sustainability variables section*). To compare high- and low-scoring companies, we construct four ESG portfolios from S&P 500 based on the ESG score provided by Refinitiv, resulting in two top portfolios and two bottom portfolios. To test the relation between the ESG score and financial performance, we conduct multiple regressions using the Carhart four-factor model and the ESG score. To further test the financial performance, we compare the risk-adjusted alphas between the top and bottom portfolios.

Friede, Busch, and Bassen (2015) report that most previous studies find a positive relation between ESG criteria and financial performance. In accordance with their research, we have reason to believe that a portfolio consisting of companies with high ESG scores performs better financially than a portfolio consisting of companies with low ESG scores. There are, however, studies suggesting the opposite. They find a negative relationship (Sahut & Pasquini-Descomps, 2015) or cannot establish a significant result (Atan, Alam, Said, & Zamri, 2018). In addition, since the ESG score should reflect sustainability, we have reason to believe that a portfolio consisting of companies with high ESG scores performs better on a range of sustainability metrics than a portfolio consisting of companies with low ESG scores. However, it is relevant to note that two reports by the OECD (OECD, 2020; Boffo & Patalano, 2020) recently found a negative relation between ESG score and environmental performance. In accordance with what we expected, we find that a portfolio consisting of companies with high

ESG scores performs better on most of the chosen sustainability metrics than a portfolio consisting of companies with low ESG scores. In addition, we find a negative relation between ESG score and financial performance, which contradicts the research by Friede et al. (2015). The results imply that ESG score screening is a sufficient method to allocate investments to companies with top sustainable performance, meaning that investors can rely on the rating when wanting to contribute to a more sustainable society. However, the results also imply that screening of ESG scores is not a sufficient strategy to target companies with positive financial performance.

Given the growing interest in sustainable investing and the divergent results in previous literature, much research is left to be done to understand how to fully assess the sustainable impact of investments. Previous research has mainly focused on the relation between ESG score and financial performance, which is continuously important as it lies in each investor's interest to make a profit. Few studies investigate the relation between ESG score and sustainable performance. As previously discussed, the financial market can be a powerful tool contributing to a positive, sustainable change. However, for the financial market to truly serve as a catalyst, the information at hand must be correctly assessing the sustainable impact. With this research, we attempt to assess if investments in companies with high ESG scores translate into positive financial *and* sustainable performance. Fund managers need to be able to strengthen that the invested capital is making an actual impact; otherwise, there is a risk of box-ticking and window dressing, which could result in decreased trust from investors. Through this research, we contribute with knowledge about ESG investing and sustainable performance, which can aid all types of investors to make informed decisions where they can rely on the investment to have a positive, sustainable impact.

2. Literature review

This section provides a review of the existing literature in the research area. First, literature regarding the link between financial performance and ESG scores, ESG disclosures, CSR scores, and CSR disclosure is presented. Second, literature related to ESG and sustainable performance, and third, literature discussing the usefulness of ESG scores in investment decisions. The ESG score is built upon the disclosure made by firms on ESG issues; hence the disclosure is of interest. Although CSR and ESG are not the same, ESG has grown from the CSR concept, and CSR can be seen as the precursor of ESG (Alva Group, 2020). CSR was initially viewed as an add-on by companies, while ESG is a more developed and integrated part of how companies do business (Alva Group, 2020). Even though ESG is growing more dominant it is still common to use the two terms interchangeably (e.g., Gillan, Koch & Starks, 2021; Fatemi, Glaum & Kaiser, 2018), which we also do in this thesis.

2.1. ESG and financial performance

The first strand of research revolves around the relationship between a company's financial performance and ESG, where the results are divergent. This area has received much attention in previous literature and is well-researched (Friede et al., 2015; Velte, 2017). Friede et al. (2015) discovered that around 90 percent of the 2,200 reviewed studies in their paper find a non-negative relationship between ESG and financial performance and that a large majority find a positive relationship. Sustainability performance is difficult to quantify, and many researchers use the ESG score retrieved from sustainability rating agencies as a proxy (Drempetic, Klein & Zwergel, 2019). In line with Friede et al. (2015), Velte (2017) points out the positive links from previous studies while also highlighting inconsistencies in the results with some negative or non-significant relations presented in the literature. One of the main reasons for the inconsistencies is the differences in the measurement of ESG performance and financial performance (Velte, 2017; Brooks & Oikonomou, 2018; Gregory, Tharyan & Whittaker, 2014). Li, Gong, Zhang, and Koh (2018) state that the differing results within the field are due to measurement concerns, data constraints, or model misspecification. In addition, Mănescu (2011) describes the possibility of a no-effect scenario when there is no difference for the abnormal return of high ESG companies compared to low ESG companies. Suppose the ESG performance of companies does not provide any information relevant for pricing. In that case, it is in line with the efficient market hypothesis that all available information is reflected in the price; see *Appendix A* for further background. The no-effect scenario could stem from an efficient valuation of ESG by the market, meaning that regardless if ESG contains relevant information or not, it will be incorporated in the price, hence not affecting the stock returns (Mănescu, 2011). In a report published by Morningstar, Wang and Sargis (2020) created an ESG factor to investigate if there exists a premium in returns as investors pay higher prices for good ESG companies compared to poor performing ones. The authors find that investors pay a slight premium for holding good ESG stocks in the North American market.

One of the primary motivations for the positive relation between ESG and firm performance is a better understanding of a firm's risk through reduced information asymmetry. Fatemi et al. (2018) research the relation between ESG activities and firm value and find that ESG disclosures have a mediating role. Further, stating that ESG disclosures in general decrease firm value, they note that increased value could be expected due to the decreased information asymmetry through more disclosures and a better understanding by investors of a firm's ESG strengths or weaknesses. On the other hand, decreasing firm value could be due to investors viewing the increased disclosures as greenwashing or window dressing. Di Giuli and Kostovetsky (2014) explain that increases in CSR rating, made through increased focus and investments, are associated with declining ROA and negative future stock returns. Suggesting that it is costly for a firm to be "good" and that the costs will not recover through increased sales; hence CSR is at the expense of firm value. Fatemi et al. (2018) find that a company with ESG concerns benefits by disclosing more to mitigate the negative effects, while a company with ESG strengths gets a lower firm valuation when increasing their disclosures by decreasing the positive effect of their strengths. Li et al. (2018) explain that ESG disclosures give additional information about the financial data provided by companies. Since ESG disclosure is a significant part of the non-financial information provided by a company, it helps to create a better understanding of its business and improves the price informativeness. The authors find a positive relationship between ESG disclosure level and firm value. They further suggest that through improving transparency and accountability and strengthening stakeholder trust, the value of a firm gets boosted.

Another motivation for the positive relation between ESG and firm performance is the lower cost of capital. Plumlee, Brown, Hayes, and Marshall (2015) find a relation between voluntary disclosures, the disclosure quality of environmental reporting, and future cash flow and cost of equity. The authors refer to traditional economic theory and motivate that a decreased cost of capital is related to increased voluntary disclosures because of the decrease in information asymmetry, enlarged investor base, and/or increased liquidity. Similarly, El Ghoul, Guedhami, Kwok, and Mishra (2011) find that firms with better CSR scores have cheaper equity financing, which is in line with their hypothesis building on the riskiness of a firm, information asymmetry, and agency problems. The cost of a firm's equity is the required rate of return, given the market's perception of its riskiness. Hence, CSR reporting provides more information to investors and affects the perceived riskiness, meaning that firms engaged in CSR should also get lower equity financing costs. Further, with more strict disclosure standards and effective corporate governance, it should lower the cost of equity through reduced information asymmetry and agency problems (El Ghoul et al., 2011). Similarly, Cheng, Ioannou, and Serafeim (2014) find that companies with better CSR performance have significantly lower capital constraints. The main reasons for the lower capital constraints are decreased information asymmetry between the company and investors due to improved reporting of CSR activities and greater transparency. Eccles, Ioannou, and Serafeim (2014) divided companies as high sustainability companies, meaning companies that voluntarily adopted sustainability policies by 1993, and low sustainability companies, meaning those that barely adopted any of them. The high sustainability companies significantly outperformed the low ones in the long term, both on the stock market and in accounting terms. Buallay (2019) further highlights aspects such as how ESG can be used as a marketing tool to get capital or as a way to indicate how companies control their business risks. Hence the higher the ESG score, the lower the business risks.

To summarize, the research within ESG and financial/firm performance is extensive. Still, as Gillan et al. (2021) discussed, there is more to discover and an ongoing debate on the subject and the underlying reasons for the links found. In the existing literature, researchers use both different variables and methods. Some studies use classical econometric models such as the CAPM, Fama French, Fama-MacBeth, or the Carhart four-factor model (e.g., Derwall, Guenster, Bauer & Koedijk, 2005; Sahut & Pasquini-Descomps, 2015; Bolton & Kacperczyk, 2020; Mănescu, 2011). Some commonly used factors are Tobin's Q, Stock price return, WACC, ROE, and ROA (e.g., Atan et al., 2018). Further, some studies conduct portfolios and compare between them or with a benchmark (e.g., Derwall et al. 2005; Verheyden, Eccles & Feiner, 2016).

2.2. ESG and sustainable performance; motives for sustainable investing

The second strand of research focuses on the SDGs and addresses the motives behind and outcome of sustainable investments. As previously mentioned, there is much research investigating the relation between ESG and financial performance. There is, however, less research investigating the sustainable impact of ESG investments. Nevertheless, the topic has raised discussion by financial experts in recent years and in 2020, OECD released two reports arguing that top ESG portfolios are not necessarily aligned with strong, sustainable performance (OECD, 2020; Boffo & Patalano, 2020). More specifically, they focused on environmental performance and raised the concern that the investors' expectations may not be met. Apart from these reports, it is not easy to find research looking at the sustainable impact of ESG investing. However, much literature investigates the motives behind sustainable investments (e.g., Jansson & Biel, 2011; Amel-Zadeh & Serafeim, 2018; Schramade, 2017), which gives a greater understanding of the desired sustainable impact of ESG investing.

In an attempt to assess the motives behind sustainable investments, Jansson and Biel (2011) compared private investors, institutional investors, and fund managers. They found that different types of investors have different motives, divided into environmental, social, and financial values. Amel-Zadeh and Serafeim (2018) elaborates on these motives by analyzing why investors use ESG information in their investment decisions. They refer to the motives as financial and impact-based. Schramade (2017) further enhanced the impact-based motive and argued that the SDGs could be used as means to make an impact. The author further argued that it is essential for companies and investors to embrace these goals to contribute to society. The motives can be boiled down into specific aspects of sustainability. Matos (2020) mentions 11 groups of sustainable issues as the most prominent factors for investors with an impact-based motive, as seen in Table 1.

Table 1. Main ESG issues (Matos, 2020)

Environmental	Social	Governance
Climate change and carbon emissions	Workforce health and safety, diversity and training	Shareholder rights
Natural resource use and energy and water management	Customer and product responsibility	Composition of boards of directors (Independence and diversity)
Pollution and waste	Community relations and charitable activities	Management compensation policy
Ecodesign and innovation		Fraud and bribery

Notes: The 11 groups of ESG issues presented by Matos (2020) divided up for each ESG pillar.

Jansson and Biel (2011) further investigate the difference in motives between private investors, institutional investors, and fund managers at investment institutions. They found that environmental concerns and the need for social justice and equality were prominent among the private and institutional investors, whereas financial performance was more important for investment institutions. The findings mentioned above align with the results from the research by Amel-Zadeh and Serafeim (2018). They show that financial motives have primarily driven the use of ESG information in investment decisions for mainstream investment organizations. There is, however, a significant group with ethical concerns as their primary motive. According to Jansson and Biel (2011), the difference in motives can be problematic, as investment institutions miscalculate what is essential for their beneficiaries. More specifically, the authors find that investment institutions underestimate the importance of environmental and social performance and the exclusion of sin companies. That is, the exclusion of companies engaging in, what some people consider, unethical or immoral industries, such as tobacco, alcohol, pornography, and weapons (Jansson & Biel, 2011). Statman and Glushkov (2009) find that investors that want to do good for society and well financially should invest in companies with high social responsibility scores but refrain from excluding sin companies. By excluding sin companies and only investing in good performing companies in terms of social responsibility, the authors find a no-effect scenario, where the returns are almost equal to conventional stocks. To extend the relation between ESG and the SDGs, Schramade (2017) links the SDGs with different sub-industries. The author argued that the SDG 8 "Decent Work and Economic Growth" applied to almost all companies with a viable business model. Hence, the conclusion was that companies could generate a good score on the variables related to decent work and economic growth without performing well on the overall ESG score; a company can perform poorly on the ESG score and yet contribute to the SDG 8 by creating jobs.

The research mentioned above relates to our study because the main drivers to incorporate sustainability factors in investment decisions are targeted. These drivers give a deeper understanding of which factors are essential to ensure that the desired result is achieved. The conclusions in previous research are twofold. One of the main drivers is the financial motive, especially in the US market (Amel-Zadeh & Serafeim, 2018), enhancing the importance of exploring the relation between ESG and financial performance. In contrast, the motive to make an impact seems more significant in Europe (Amel-Zadeh & Serafeim, 2018), and the UN SDGs are mentioned as means to make an impact (Schramade, 2017). Factors such as CO2 emission, gender equality, and overall contribution to society seem to be the most prominent ones when discussing the impact.

2.3. The usefulness of ESG scores

The third strand of research relates to ESG scores and the discussion regarding their usefulness. There are several providers of ESG scores, and there is some concern regarding the divergence between the ratings and the usefulness of the ESG scores (Berg et al., 2020b; Drempetic et al., 2019; Gibson, Krueger & Schmidt, 2019). Berg et al. (2020b) find that the difference in rating from different rating agencies depends on the scope of categories, measurements of the categories, and weighting. Hence, given the divergence between rating agencies, it can be difficult for investors, or other users of ESG scores, to understand the underlying performance that results in a particular ESG score. Drempetic et al. (2019) further highlight the question of whether the ESG score provides investors with the information needed to make well-informed decisions in line with their beliefs. The authors even suggest that some ESG scores do not provide the information required by investors and researchers. Schramade (2017) claims that ESG ratings have limitations and encourage investors to move beyond the ratings and conduct their own research on topics important to them to get valuable insights.

Further, Drempetic et al. (2019) find indications of a firm-size bias, meaning that larger firms with more resources tend to get a higher ESG score due to the measurement of the score. Gibson et al. (2019) provide evidence in their study that the divergence in scores is higher for the largest firms on the S&P 500 and companies that do not have a credit rating. Not only the size but also the legal origin seem to play a part in the rating of a company. Liang and Renneboog (2017) find a strong correlation between a company's CSR rating and its country's legal origin. More specifically, companies originating from common law countries have lower CSR ratings than companies originating from civil law countries. Similarly, Baldini, Maso, Liberatore, Mazzi, and Terzani (2018) find that a companies' ESG disclosure practice is significantly affected by country-level characteristics.

As mentioned in the introduction, the number of firms reporting on ESG issues and receiving an ESG score has increased. A key driver for the increased attention towards ESG when making investment decisions is the demand for it by institutional and individual investors (Mănescu, 2011). In Europe, it has been mandatory for large publicly listed companies to report on non-financial performance since 2018. Currently, the directive is under review to further strengthen the base for sustainable investments (European Commission, 2021). On the other hand, it is not mandatory in the US, although there is an ongoing debate on the subject, as Fleming and Ledbetter (2020) discussed in a white paper. The quality of a report is often evaluated based on the usefulness for a reader (Runesson, Samani & Marton, 2018). The difficulty with sustainability reporting is that there are many stakeholders of a firm, and hence, it is difficult for firms to provide high-quality reporting for all the users. Another difficulty is the comparability between firms on non-financial metrics, which is a highlighted aspect in many non-financial reporting directives (D'Aquila, 2018; Rupley, Brown & Marshall, 2017). Depending on industry and firm-specific factors companies need to determine material features of their business (D'Aquila, 2018; Rupley et al., 2017). This can be problematic as it opens up for interpretation.

Further, an important aspect to point out regarding non-financial reporting is that there are several different, and sometimes competing, frameworks that companies can choose to apply (Jensen & Berg, 2012; Rupley et al., 2017; D'Aquila, 2018). With the complexity of sustainability reporting in mind, it is clearly a difficult task to try and use all the information provided by companies on non-financial performance and landing in one quantifiable number. To summarize, ESG reporting and ESG ratings are difficult areas with many dimensions. It can be challenging to quantify and compare non-financial performance reporting. Further, summarizing all the information into one number is even more difficult. The ESG rating providers are experts in the field; however, divergence occurs, and there is still not one easy and complete way to measure the complexity of a firm's actions into a number or a grading.

3. Hypotheses

Our main objective is to evaluate if ESG investing contributes to a more sustainable society. More specifically, to answer whether screening of ESG scores is a sufficient method to allocate investments to companies with top sustainable performance. There are, of course, several ways of measuring the sustainable outcome. However, to make this feasible for our research, 10 sustainability metrics are chosen, as described in the *Sustainability variables section.*

The relation between ESG score and financial performance is evaluated to explore the main objective further. According to portfolio theory (*Appendix B),* investors seek to maximize return given a certain level of risk. Hence, to attract capital, it is relevant to investigate the financial performance of companies with high ESG scores.

One main hypothesis is formulated to explore the relation between ESG score and sustainable performance, and two sub-hypotheses to explore the relation between ESG score and financial performance. Our approach to testing the hypotheses is described in the *Methodology section.*

3.1. Main hypothesis

The main hypothesis addresses whether screening of ESG scores is a sufficient method to allocate investments to companies with top sustainable performance over the studied period. Hence, hypothesis I test whether top ESG portfolios perform differently than bottom ESG portfolios on the selected sustainability metrics.

Hypothesis I

H0: *Top ESG portfolios perform similar to bottom ESG portfolios measured on a range of sustainability metrics*

H1*: Top ESG portfolios perform differently from bottom ESG portfolios measured on a range of sustainability metrics*

3.2. Sub-hypotheses

The relation between ESG score and financial performance is tested through the sub-hypotheses. We also test if either of the top portfolios performs differently financially than the bottom portfolios. The relation is tested using the Carhart four-factor model with the ESG score, ESGC score, and an ESG factor added as an extra variable. The ESGC score, also called the combined ESG score, is an extension of the ESG score, which also considers news controversies that materially impact the companies. See *Appendix C* and the *Methodology section* for further description of the model and the *Data section* for further description of the ESG factor. The performance is also tested by comparing the risk-adjusted alphas through the Carhart four-factor model using a difference portfolio.

Hypothesis II

H0*: There is no relation between the ESG score and financial performance*

H1: *There is a relation between the ESG score and financial performance*

Hypothesis III

H0*: The risk-adjusted alphas are similar between top portfolios and bottom portfolios*

H1: *The risk-adjusted alphas are different between top portfolios and bottom portfolios*

4. Data

In this section follows a description of the data and reasoning behind the chosen data. In short, data was collected from three sources: Refinitiv Eikon, S&P Capital IQ, and Kenneth R. French data library. The data was merged based on trading days to carry out the analysis. The result is a time-series analysis of panel data of 480 listed US equities, ranging from the start of 2015 to the end of 2019. The stock data and the factors from Kenneth R. French data library were collected daily, screened on trading days. The sustainability data was collected yearly, as it is reported.

All handling of the data and the statistical analysis is conducted in Stata version 16 and Microsoft Excel.

4.1. Time period

As stated in the introduction, ESG became a buzzword in the financial industry following the introduction of the UN SDGs in 2015 (PwC, 2020). With this background, 2015 is chosen as the starting year of our study. The lack of data during the years before 2015 further strengthened this decision. According to D'Aquila (2018), ESG reporting has increased dramatically in recent years; less than 20 percent of the S&P 500 companies reported sustainability data in 2011, which increased to above 80 percent in 2015.

Further restrictions of the period are made due to the COVID-19 pandemic. 2020 was a year marked by the pandemic, which to a large extent affected the financial markets (Donthu & Gustafsson, 2020). It would be interesting to investigate the impact of COVID-19 on ESG score, ESG performance, and the financial market. However, the effects of COVID-19 are out of the scope of this study, thus 2020 has been excluded from the studied period.

4.2. Sample

ESG reporting is not yet mandatory in the US, as it is in Europe (European Commission, 2021). Although Socially Responsible Investments (SRI) are more prominent in Europe, the US is experiencing the largest growth in SRI invested assets (Verheyden et al., 2016). Collins and Sullivan (2020) predict that half of all the managed assets in the US could be ESG mandated by 2025. Hence, we find it interesting to investigate how ESG investments affect sustainable development in the US, which is the reason for choosing S&P 500 as the sample.

First, the sample was screened based on companies in S&P 500, which had an ESG score in 2015 (n=491). Further screening was made based on ESG score and stock price availability, and 11 companies were removed from the sample. Lack of available stock price could be due to an introduction to the stock market during 2015 or because of a split or merger. As shown in Table 2, the final sample consists of 480 companies.

Table 2. Sample screening

Screening for	Missing	Total
S&P 500		505
ESG score 2015	−14	491
ESG score available	−4	487
Stock price available for the whole period	−7	480
Full sample		**480**

Notes: The screening process and removal of companies based on data availability.

4.3. ESG data

The ESG data is retrieved from Refinitiv Eikon during the spring of 2021. Refintiv Eikons' ESG score is widely used in the financial literature, and Berg, Fabisik, and Sautner (2020a) refer to Refinitiv as one of the leading providers of ESG data. The database covers over 70 percent of the global market cap and offers one of the most comprehensive platforms for ESG rating (Refinitiv, 2021). The metric is designed to provide transparent, comparable, and accurate data on companies' relative ESG performance, effectiveness, and commitment based on company-reported data. The ESG score is divided into 10 main categories, further evaluating over 500 different aspects of sustainability (Refinitiv, 2021). The ESG score, the ESGC score, and a subset of the different sustainability aspects are used to conduct the analysis (see Sustainability variables for further description). The purpose of the ESGC is to adjust the ESG performance based on negative news stories (Refinitiv, 2021). The ESGC score and the ESG score of a company are equal if the company is not involved in any ESG controversies.

When Refinitiv computes the ESG score, the different categories are weighted proportionally based on the number of measures within each category, which generates the ESG score on a scale ranging from 0 to 100. The current weights for the 10 categories are defined in Table 3.

Table 3. Refinitiv category weights

Pillar	Category	Weight*
Environmental	Emissions	15%
	Resource use	15%
	Innovation	13%
Social	Community	9%
	Human rights	5%
	Product responsibility	4%
	Workforce	13%
Governance	Shareholders	5%
	CSR strategy	3%
	Management	17%

Notes: The weights for each category used in the calculation of the ESG score provided by Refinitiv. *Add up to 99 % due to rounded numbers.

Source: Refinitiv (2021). For further understanding of the construction of the ESG score, visit Refinitiv.com.

4.4. Sustainability variables

In addition to the ESG and ESGC scores, some sustainability variables are used to discuss the sustainable impact of the portfolios. These variables are retrieved from Refinitiv Eikon. To ensure the relevance of the factors, they are chosen based on three criteria, which are as follows:

1. Applicable to the United Nations' Sustainable Development Goals.
2. Variables that have been considered essential for each pillar in previous literature.
3. Variables available through Refinitiv Eikon for a large majority of the companies in the sample.

According to a report released by OECD (2020), there is a core set of environmental factors that negatively affects the environment. These are CO2 emissions, total waste production, total energy use, and total water withdrawals. Matos (2020) further stresses the importance of these factors, and adds a few factors in the Social and Governance pillars, as shown in Table 1 in the *Literature review section*. These factors are further enhanced by Amel-Zadeh and Serafeim (2018), who present a few examples on environmental, social, and governance data. They emphasize carbon emission, water consumption, and waste generation when discussing environmental data. Employee makeup, product information, and customer-related information when discussing social data and political lobbying, anti-corruption programs, and board diversity when discussing governance data.

As previously mentioned, the ESG score is divided into 10 main categories, and within each of these categories, there is a subset of sustainability measurements. Based on the SDGs, the metrics discussed in previous literature, and the availability on Refinitiv Eikon, we have chosen the variables shown in Table 4.

Table 4. Chosen sustainability variables

Pillar	Category	Sustainability variables	Variable name	Form
Environmental	Emission	Total CO2 Emissions / Million in Revenue $ (scope 1+2)	CO2R	Numerical, tonnes/$M
		Environmental partnership	envpart	Binary, Y/N
	Resource use	Environmental supply chain partnership termination	envpartterm	Binary, Y/N
	Innovation	Environmental products	envprod	Binary, Y/N
Social	Product responsibility	Policy responsible marketing	respmarket	Binary, Y/N
		Policy customer health and safety	customer	Binary, Y/N
	Workforce score	Net employment creation	employment	Numerical, percentage
Governance	Management	Policy board independence	boardind	Binary, Y/N
		Board gender diversity	gendiv	Numerical, percentage
	Shareholder score	Shareholder rights policy	shareright	Binary, Y/N

Notes: The sustainability variables (which are further explained in *Appendix D*) with corresponding variable names and the form it is measured in, divided into different categories, which in turn belong to one of the ESG pillars.

4.5. Financial variable

The financial variable used in this study, also referred to as financial performance, is computed as the daily return in stock price for each company in the sample. As provided by S&P Capital IQ, the adjusted closing price is used to compute the daily return (see Equation 1). Adjusted closing price means that the price is adjusted for dividends and splits.

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}}$$

Equation 1

R_t = The daily stock return for the stock, at time t

P_t = The stock price in time t

Source: Brooks (2019)

4.6. Kenneth R. French data library

The daily Fama French three-factors, SMB, HML and Mkt-RF, the Momentum, and the risk-free rate, are retrieved from the Kenneth R. French data library. The Fama French factors are created using six value-weight portfolios constructed based on size and book-to-market value (French, 2021). The one-month Treasury bill from Ibbotson Associates is used as the risk-free rate. Below is a short description of the variables; see *Appendix E* for equations.

The SMB, Small Minus Big, represents the size premium. It is calculated as the average return of three small portfolios minus the average return of three big portfolios. HML, High Minus Low, represents the value premium. It is calculated as the average return of two value portfolios minus the average return of two growth portfolios. Mkt-RF represents the market's excess return, computed as the value-weighted return for stocks incorporated and listed on the US market minus the risk-free rate. Mom, the daily Momentum, is calculated using six value-weight portfolios based on size and prior return. It is further computed as the average return on two high prior return portfolios minus the average return on two low prior return portfolios.

4.7. ESG factor

In addition to the ESG score and ESGC score, we create an ESG factor to evaluate the relation between ESG score and financial performance. The factor is created using a similar approach to Wang and Sargis (2020) and how Fama and French compute their SMB and HML factors.

The factor is created by mimicking the results one would get by buying companies with high ESG scores and short-selling companies with low ESG scores. Similar to how French (2021) computes the Momentum factor, the companies are sorted on ESG score and divided into two portfolios; the top 30 percent and the bottom 30 percent. The portfolios are rebalanced yearly to include the best-performing and worst-performing companies in terms of ESG scores each year. The daily ESG factor is created by taking the average daily returns for the best-performing companies minus the average daily return for the worst-performing, resulting in a daily ESG factor.

$$ESG\ factor_t = R_{top\ 30\ \%\ portfolio,t} - R_{bottom\ 30\%\ portfolio,t}$$

Equation 2

4.8. Descriptive statistics

Table 5 presents descriptive statistics. As mentioned in the *Sample section*, companies with missing data on ESG score or adjusted closing price for the entire period are removed, resulting in 603,840 observations (N) for each variable. As presented in the table, some variables have fewer observations due to missing data on some of the sustainability variables. The variables are kept as missing, without a proxy, to discuss data availability for companies with different ESG scores. In the *Sustainability variables section*, Table 4 presents that some variables are binary, only taking values 0 or 1, which is also apparent when looking at the minimum and maximum values in Table 5 below. When the mean and the median are far from each other, it entails that the data is not normally distributed but somewhat skewed (Salkind, 2010), as seen in the *CO2R* variable. The skewness in *CO2R* could be because different industries have different emission levels, and the companies are not equally distributed between the industries (as presented in Chart 1). Table 5 also shows that there are some large outliers. As *CO2R* is not used in a linear model, we chose not to remove or adjust large outliers.

Table 5. Descriptive statistics

Variable	N	Mean	Median	Std. Dev.	min	max
R-RF	603840	-.003	-.003	.017	-.349	.523
Mkt-RF	603840	.043	.05	.866	-4.03	5.06
SMB	603840	-.008	-.01	.498	-1.67	2.51
HML	603840	-.017	-.05	.554	-1.86	3.07
Mom	603840	.015	.055	.792	-3.81	3.63
ESG	603840	57.93	60.22	17.551	2.48	93.13
ESGC	603840	54.252	54.6	16.84	2.48	92.62
ESGF30	603840	0	0	.003	-.009	.013
gendiv	596282	22.539	22.22	9.121	0	62.5
CO2R	387679	407.197	41.53	1071.868	0	7732.21
employment	600318	.063	.03	.243	-.83	6.23
customer	597037	.363	0	.481	0	1
envpart	597037	.555	1	.497	0	1
envprod	597037	.531	1	.499	0	1
envpartterm	597037	.316	0	.465	0	1
boardind	597037	.976	1	.152	0	1
respmarket	597037	.05	0	.218	0	1
shareright	597037	.997	1	.05	0	1

Notes: The descriptive statistics of the variables used for all three hypotheses where N is the number of observations.

In Table 6, the pairwise correlation of the variables used in the regressions is presented. These are the variables used to address hypotheses II and III. Most of the variables have some significant correlation with other variables. For example, *ESG* and *ESGC* have a strong significant correlation of 0.841, which is reasonable since the calculation of the scores is very similar. However, this correlation is not a problem since the two variables are not used in the same regression; *ESGC* is only used instead of *ESG* as a robustness check. Other variables with relatively high correlation are *Mkt-RF* and the dependent variable *R-RF (Return-Risk free rate)*. Further, the *MOM* and *HML* variables from the Carhart model have a relatively high correlation (-0.461). However, we do not see it as problematic since the variables are retrieved from the Kenneth R. French data library, and the models are well-referenced econometric models. Like the study by Wang and Sargis (2020), the *ESGF30* has a relatively high negative correlation with the size factor *SMB* (-0.437). The high correlation indicates that firms with higher ESG scores tend to be larger companies, which is in line with previous literature (e.g., Drempetic et al., 2019).

Table 6. Pairwise correlation matrix

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) R-RF	1.000							
(2) Mkt-RF	0.496***	1.000						
(3) SMB	0.086***	0.140***	1.000					
(4) HML	-0.016***	-0.086***	-0.084***	1.000				
(5) Mom	-0.078***	-0.101***	-0.187***	-0.461***	1.000			
(6) ESG	-0.033***	0.004***	-0.001	-0.003**	-0.002	1.000		
(7) ESGC	-0.022***	0.002*	-0.001	-0.003**	0.000	0.841***	1.000	
(8) ESGF30	-0.137***	-0.274***	-0.437***	0.319***	-0.131***	0.002*	0.001	1.000

*Not*es: The pairwise correlations for all the variables used when testing hypotheses II and III. The stars denotes the p-value; *** p<0.01, ** p<0.05, * p<0.1

4.9. Portfolio composition

Similar to how Derwall et al. (2005) created top and bottom portfolios based on eco-efficiency, we create the top and bottom portfolios based on the ESG score. Hence, the portfolios are created by selecting the top and bottom scoring companies. The portfolios are binary variables, taking value 1 if the company is in the top portfolio and 0 if the company is in the bottom portfolio, otherwise left blank. Since we are interested in exploring the difference in performance between the top and bottom portfolios, a difference portfolio is also created. The difference portfolio is constructed by subtracting the average daily return for the bottom portfolio from the top portfolio's average daily return. Derwall et al. (2005) have previously used this technique to compare a top and bottom portfolio.

We create two top portfolios, one including the top 10 percent of the best ESG scoring companies and one including the top 30 percent. These portfolios are compared to the bottom 10 percent and the bottom 30 percent in terms of ESG scores to test if the top-scoring companies perform significantly better. Since the ESG score is updated annually, the portfolios are rebalanced yearly. Thus, the composition changes depending on what score the company received each year from 2015 to 2019. Meaning that some companies are removed, and others added each year. On average, 25 percent of the companies in the top and bottom 10 percent portfolios are replaced each year, while 15 percent of the companies in the top and bottom 30 percent portfolios are replaced each year.

Table 7 shows the limits for the ESG scores used when yearly rebalancing the portfolios. Note that the breaking point for the ESG score for all portfolios is constantly increasing. There could be several reasons for the increased limit, e.g., that the companies have improved their ESG work or increased their ESG disclosure.

Table 7. ESG breaking point for portfolio creation

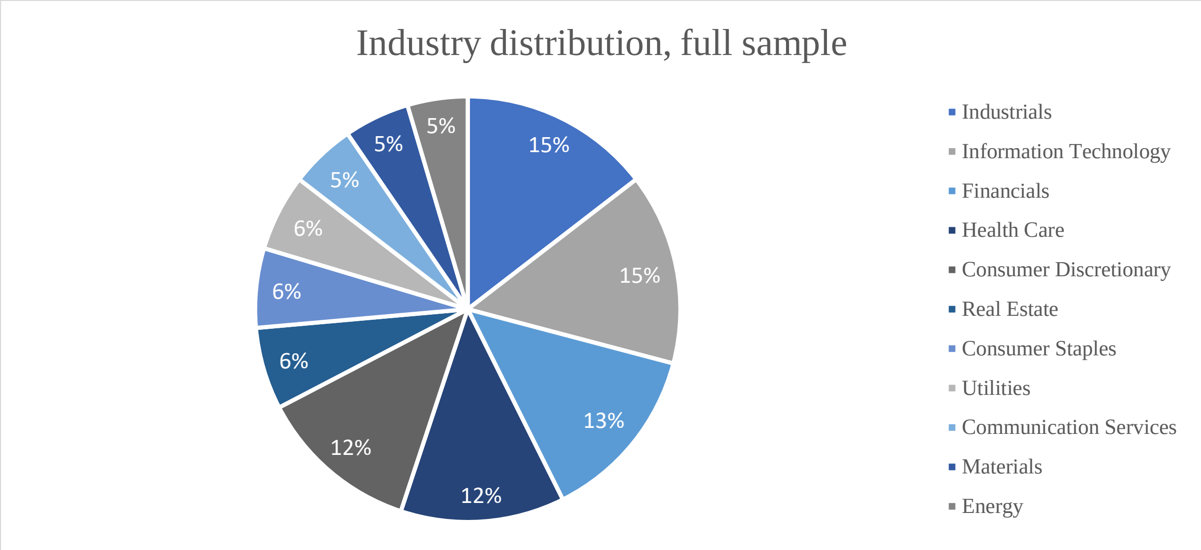
Portfolios, ESG score limit	2015	2016	2017	2018	2019
Portfolio Top 10% - 48 companies	77.3	78.6	78.98	79.74	81.37
Portfolio Bottom 10% - 48 companies	27.12	32.62	32.95	36.37	39
Portfolio Top 30% - 144 companies	64.96	68.66	69.32	70.35	71.9
Portfolio Bottom 30% - 144 companies	42.87	46.17	49.2	51.2	55.3

Notes: The limits of the ESG scores used for creating the portfolios, balanced yearly.

4.10. Industry distribution

The GICS, Global Industry Classification Standard, is used for the industry classification of our sample. GICS classifies companies quantitatively and qualitatively, and revenues are a key factor in determining the main business (MSCI, 2021). As presented in the *Sample section*, the sample consists of 480 companies. These companies are classified into 11 different sectors, as shown in Chart 1. Note; in the full sample, more than half of the companies are grouped in four out of the 11 industries. The industry distribution of the top and bottom portfolios based on the ESG scores is found in *Appendix F*.

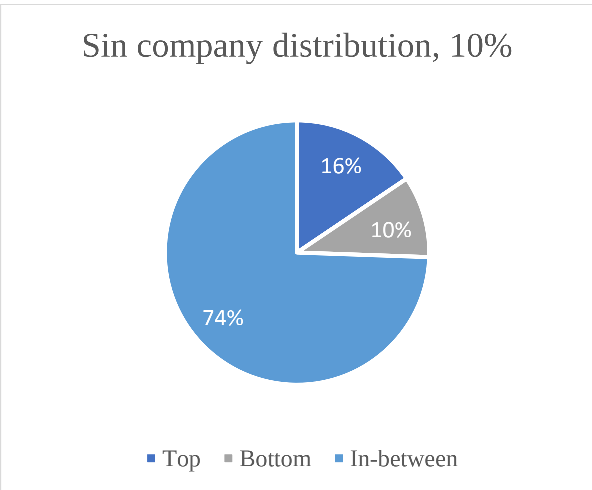
Chart 1. Industry distribution, full sample

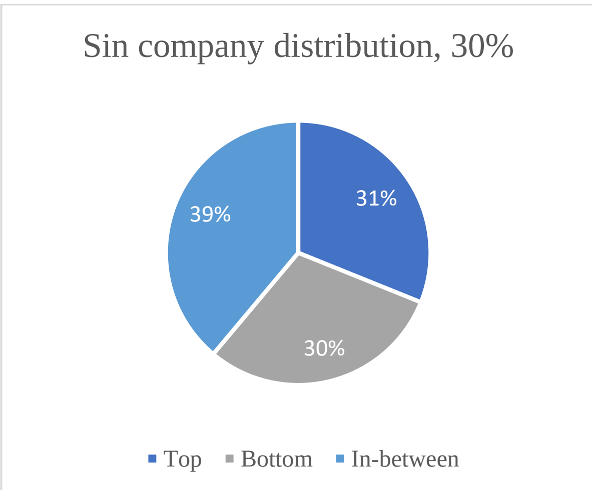


Notes: The full sample (480 companies) divided in the 11 GICS sectors.

The companies can be further grouped into industries and sub-industries to get a more specific classification. As discussed in the *Literature review*, it is important for some investors that the portfolio excludes sin companies. To explore the sin companies in the sample, Chart 2 describes how the sin companies are distributed in the portfolios. Sin companies are referred to as companies engaging in tobacco, alcohol, gambling, and weapons (Jansson & Biel, 2011), equivalent to the GICS sub-industries; aerospace and defense, casinos and gaming, distillers and vintners, brewers, and tobacco. As depicted in the charts, 16 percent of the sin companies are in the top 10 percent portfolio, 10 percent are in the bottom 10 percent portfolio, 31 percent are in the top 30 percent portfolio, and 30 percent are in the bottom 30 percent portfolio.

Chart 2. Sin company distribution





Notes: The average distribution of sin companies in the different portfolios for the studied period.

5. Methodology

t-Tests are used to explore our main hypothesis and test if there is a significant difference in sustainable performance between the top and bottom portfolios. The sub-hypotheses are further explored by running General Least Squares (GLS) regressions. In addition, regressions using the CAPM and the Fama French three-factor model are conducted to test the robustness of the sub-hypotheses. See *Appendix G* and *Appendix H* for further descriptions of the models. The reasoning for the chosen methodology is further developed under each section.

5.1. t-Test

When evaluating the main hypothesis, the sustainable performance of the different portfolios is of interest. In order to test if companies with high ESG scores perform better on the chosen sustainability variables, the mean is tested with a t-test. A t-test is used to determine if the mean of two samples is significantly different from each other (Salkind, 2010). The null hypothesis is that the two means are equal, tested through a two-sided t-test. The t-test can also be one-sided, determining if one sample has a higher or lower mean. Since the direction is of interest, i.e., assessing if the top portfolio performs better or worse than the bottom portfolio, both a one- and two-sided t-test is used. Through the one-sided t-test, we test which of the portfolios performs better in two steps. First, it tests if the mean of the bottom portfolio is greater than the mean of the top portfolio. Second, it tests if the mean of the top portfolio is greater than the mean of the bottom portfolio.

5.2. Econometric model

As previously mentioned, the primary econometric model to address hypotheses II and III is the Carhart four-factor model. The model is a development from the Fama French three-factor model (1993) and includes an additional momentum factor; see *Appendix C* for background and specification of the model. For example, Mănescu (2011) applies the four-factor model regarding ESG and risk-adjusted stock returns, and so does Derwall et al. (2005) in their research regarding eco-efficiency and stock returns.

Hypothesis II, whether there is a relation between ESG score and financial performance, is tested through a GLS-regression using the Carhart four-factor model and adding an ESG variable. A similar approach is used by Mănescu (2011), where the four-factor model is extended with selected ESG variables and the ESG score. We use three different ESG variables; the yearly ESG score, the yearly ESGC score, and the daily ESG factor described in *section 4.7*. See Equation 3, Equation 4, and Equation 5 below for the tested models.

$$r_{i,t} - rf_t = \alpha_{i,t} + \beta_{1,i}(r_{mkt,t} - rf_t) + \beta_{2,i}SMB_t + \beta_{3,i}HML_t + \beta_{4,i}MOM_t + \beta_{5,i}ESG\ score_{i,t} + \varepsilon_{i,t}$$

Equation 3

$$r_{i,t} - rf_t = \alpha_{i,t} + \beta_{1,i}(r_{mkt,t} - rf_t) + \beta_{2,i}SMB_t + \beta_{3,i}HML_t + \beta_{4,i}MOM_t + \beta_{5,i}ESGC\ score_{i,t} + \varepsilon_{i,t}$$

Equation 4

$$r_{i,t} - rf_t = \alpha_{i,t} + \beta_{1,i}(r_{mkt,t} - rf_t) + \beta_{2,i}SMB_t + \beta_{3,i}HML_t + \beta_{4,i}MOM_t + \beta_{5,i}ESG\ factor_{i,t} + \varepsilon_{i,t}$$

Equation 5

A significant positive coefficient of ESG would indicate that a higher ESG score leads to increased financial performance, while a negative coefficient would indicate the opposite. Hence, if β_5 is significant, it would support the hypothesis that there is a relation between ESG score, ESGC score, or ESG factor and financial performance.

Hypothesis III, whether the four-factor alphas are different between the top and bottom portfolios, is also tested through separate GLS-regressions of the portfolios using the Carhart four-factor model. That is, whether the portfolios have a different abnormal return compared to each other. In CAPM, the alpha is called Jensen's alpha, and it measures whether the portfolio has out- or underperformed what is expected given its level of market risk (Brooks, 2019). In the Fama French three-factor model, it is called the three-factor alpha, and we are interested in the four-factor alpha, describing the daily abnormal return of the portfolio, controlling for the Carhart factors. See Equation 6, the Carhart four-factor model. Similar to Derwall et al. (2005), a difference portfolio is used to compare the alphas and test if one portfolio performs significantly better than the other.

$$r_{i,t} - rf_t = \alpha_{i,t} + \beta_{1,i}(r_{mkt,t} - rf_t) + \beta_{2,i}SMB_t + \beta_{3,i}HML_t + \beta_{4,i}MOM_t + \varepsilon_{i,t}$$

Equation 6

5.3. Fixed-effects model

In financial research, it is common to employ one of the following panel estimator approaches; fixed-effects model or random-effects model (Brooks, 2019). The random-effects model is commonly used when the sample is randomly selected from the population. In contrast, the fixed-effects model is more appropriate when the entities in the sample constitute the entire population (e.g., all of the stocks on a particular exchange). According to Brooks (2019.), the random-effects model generates a more efficient estimation than the fixed-effects approach, but it has some significant drawbacks. The drawback lies in the fact that the approach is only valid when the error term is uncorrelated with all explanatory variables. Due to our sample selection process, where the sample was not randomly selected, we argue that the fixed-effects model is preferred. In addition, due to the complexity of the ESG score, we have reason to believe that the score (i.e., our explanatory variable) can be correlated with the error term. The correlation can be tested with a Hausman test, which was, among others, done by Atan et al. (2018). Due to the scope of this thesis, we instead assume that there can be a correlation between the error term and the explanatory variable and use the fixed-effects model, still being aware that this model is not as efficient as the random-effects model.

5.4. General Least Squared, GLS

Since we are handling panel data, a suitable method to handle time-series and cross-sectional elements is needed. One way of doing so is to estimate a pooled regression and then conduct Ordinary Least Squares (OLS) regressions (Brooks, 2019). This method does, however, have limitations. Another way to estimate and test fixed-effects panel data is through GLS regression, with the command xtgls in Stata (Blackwell III, 2005). This method is strengthened by Brooks (2019), who refers to GLS regressions as a more commonly used technique when handling panel data. This method produces a more efficient estimator than the OLS estimator, and it is estimated with higher precision. The GLS method is further applicable when the number of time series observations is greater than the total number of observed entities, which is valid for our data. With this background, we argue that GLS regression is a suitable method to evaluate the second and third hypotheses.

To summarize, to use an OLS regression, a set of assumptions needs to be satisfied (see *Appendix I*). It can be tricky to satisfy the assumptions when handling panel data (Brooks, 2019). Instead, we use GLS regressions, a more flexible method as it can be used to relax one or more of the OLS assumptions (Brooks, 2019).

5.5. Robustness testing

One reason for the divergent results in previous research on ESG and financial performance is that different econometric models are used (Brooks & Oikonomou, 2018). As previously mentioned, when testing the sub-hypotheses, the primary econometric model is the Carhart four-factor model. In order to test the robustness of the results, the regressions are also conducted using two other econometric models, the Fama French three-factor model, and the CAPM; see *Appendix G* and *Appendix H* for further explanation. The results from these models are then compared with the primary model to determine whether the results depend on the used model.

5.6. Endogeneity

Endogeneity relates to the error term assumption, which addresses the problem of a correlation between an explanatory variable and the error term (Salkind, 2010). There can be multiple sources for endogeneity; when measurement issues cause that the actual value is not observed, when a variable that affects the dependent variable is missing from the regression, or when there is a feedback loop between the dependent and independent variables. It is hard to capture all explanatory variables in the regression; according to Brooks (2019), it is rare that the assumption regarding uncorrelated error term holds in practice. It is, however, possible to handle some of the potential bias due to omitted variables. One way to handle the potential correlation between the explanatory variables and the error term is to run GLS regressions as it transforms the model so that the error terms become uncorrelated (Brooks, 2019). Another way is to use a fixed-effect model since it does not require the explanatory variables to be uncorrelated with the error term (Brooks, 2019).

Even though this thesis uses methods that loosen the OLS assumptions and handle the correlation between the explanatory variables and the error term, it is important to understand that endogeneity may still exist. In this research, endogeneity could occur if another aspect drives both the ESG score and the stock return, e.g., if companies with strong finances are more likely to invest more into sustainability actions and hence receive a higher ESG score which also affects the stock return.

6. Results and analysis

6.1. Main hypothesis

Table 8 and Table 9 present the t-test for the different sustainability metrics for the top and bottom 10 and 30 percent portfolios. As shown in the tables, there is a difference in the mean between the top and bottom portfolios on all sustainability metrics. Hence, we can reject the null hypothesis at a 99 percent confidence level, suggesting that the means differ for all the chosen sustainability metrics. A larger mean denotes better performance for all variables except *CO2R*. For *CO2R*, a lower mean denotes less CO2 emissions to revenue, which is considered the better performance. Hence, a difference less than zero on all variables except for *CO2R* implies that the top portfolio performs better than the bottom portfolios. For *CO2R*, a difference greater than zero implies that the top portfolio performs better than the bottom portfolios.

Table 8. t-Tests top (1) and bottom (0) 10%

10% portfolio								
					diff	Ha: diff <0	Ha: diff != 0	Ha: diff > 0
Variable	Obs 0	Obs 1	Mean 0	Mean 1	mean(0) -mean(1)	Pr(T < t)	Pr(|T| > |t|)	Pr(T > t)
gendiv	59628	60132	16.573914	27.147171	-10.573257	0,000	0,000	1,000
CO2R	4779	59628	813.31752	267.62559	545.69192	1,000	0,000	0,000
employment	60132	60384	0.12743132	0.04738225	0.08004906	1,000	0,000	0,000
customer	59628	60132	0.09277521	0.68614049	-0.59336528	0,000	0,000	1,000
envpart	59628	60132	0.14342255	0.84933147	-0.70590892	0,000	0,000	1,000
envprod	59628	60132	0.10121084	0.7656988	-0.66448796	0,000	0,000	1,000
envpartterm	59628	60132	0.07593748	0.54385352	-0.46791604	0,000	0,000	1,000
boardind	59628	60132	0.89449587	0.98327014	-0.08877426	0,000	0,000	1,000
respmarket	59628	60132	0.02953311	0.12968137	-0.10014826	0,000	0,000	1,000
shareright	59628	60132	0.97467633	1,000	-0.02532367	0,000	0,000	1,000

Notes: t-Tests for the chosen sustainability metrics for the 10 % portfolios, where the top is denoted 1, and the bottom is denoted 0.

Table 9. t-Tests top (1) and bottom (0) 30%

30% portfolio								
					diff	Ha: diff < 0	Ha: diff != 0	Ha: diff >0
Variable	Obs 0	Obs 1	Mean 0	Mean 1	mean(0) - mean(1)	Pr(T < t)	Pr(|T| > |t|)	Pr(T > t)
gendiv	179640	178381	19.251053	25.474127	-6.2230737	0.0000	0.0000	1.0000
CO2R	44026	176366	428.6053	354.62037	73.984931	1.0000	0.0000	0.0000
employment	180146	180900	.10058897	.04563405	.05495491	1.0000	0.0000	0.0000
customer	179640	179136	.13583278	.61519181	-.47935903	0.0000	0.0000	1.0000
envpart	179640	179136	.22967045	.86093806	-.63126761	0.0000	0.0000	1.0000
envprod	179640	179136	.27308506	.7191296	-.44604454	0.0000	0.0000	1.0000
envpartterm	179640	179136	.14561345	.47752546	-.33191201	0.0000	0.0000	1.0000
boardind	179640	179136	.9495658	.98315247	-.03358667	0.0000	0.0000	1.0000
respmarket	179640	179136	.01680583	.09550844	-.07870261	0.0000	0.0000	1.0000
shareright	179640	179136	.9915943	1	-.0084057	0.0000	0.0000	1.0000

Notes: t-Tests for the chosen sustainability metrics for the 30 % portfolios, where the top is denoted 1, and the bottom is denoted 0.

The results support the alternative hypothesis that the top portfolios perform better for all chosen sustainability metrics at a one percent significance level, except for net employment creation. We find support for the alternative hypothesis that the top portfolios perform worse on net employment creation than the bottom portfolios at a one percent significance level.

To summarize, both top portfolios perform better on most sustainability metrics, which align with our expectations. Although, it is important to note that we expected the top portfolios to perform better on all sustainability metrics, which they did not. However, the results could be explained with the literature by Schramade (2017); a company can generate a good score on decent work and economic growth (SDG 8), in this case, net employment creation, without performing well on the overall ESG score. The result from the t-test confirms the research by Schramade (2017) that companies can perform better on net employment creation while worse on the other sustainability metrics. In relation to our research question and the main objective, we argue that screening of ESG scores is a sufficient method to allocate investments to companies with top sustainable performance and hence contribute to a more sustainable society. Even though the bottom portfolio performed better on net employment creation, it is important to keep in mind that the top portfolios still had a positive mean, meaning that the top portfolios still contribute to SDG 8.

In relation to previous research investigating the motives behind ESG investing (e.g., Jansson & Biel, 2011; Amel-Zadeh & Serafeim, 2018; Schramade, 2017), our study provides evidence that ESG investing can be considered a suitable method for investors with an impact-based motive, as the top ESG portfolios perform better than the bottom ESG portfolios on all except one of the chosen sustainability metrics. In addition, the results related to the first hypothesis contradict the OECD reports (OECD, 2020; Boffo & Patalano, 2020), as we find that top ESG portfolios, on average, have a stronger environmental performance than bottom ESG portfolios. However, it is relevant to note the distribution of the missing variables. As mentioned under *Descriptive statistics*, there are some missing sustainability variables, for which we did not provide a proxy. As seen in Table 8 and Table 9, the number of observations differs between the top and bottom portfolios, and the most noticeable difference is for *CO2R*. This difference makes it harder to say that the top portfolio performs better on CO2 emissions, even though the t-test provides evidence for it. Instead, a question as to why the bottom portfolio does not report on CO2 emission arises. This question remains unanswered, as an exploration of the reason behind the missing values is outside of the scope of this thesis. However, possible explanations could be that companies in the bottom portfolio do not have any CO2 emissions they find worth mentioning in their reporting or perform poorly and therefore choose not to report on it.

6.2. Sub-hypotheses

Table 10 presents the results to address hypothesis II regarding the relationship between ESG score and financial performance. The first regression, (1), is the Carhart four-factor model, with the ESG score added as an extra independent variable. The ESG score coefficient is negative and significant at a one percent level. Further, in regression (2) and (3), the coefficients for the ESGC score and the ESG factor are also significant at a one percent level and negative, which is a part of the robustness check. A significant coefficient means that the null hypothesis can be rejected at a 99 percent confidence level, suggesting a relation between ESG score and financial performance. The negative coefficient indicates that a one-unit change in ESG would affect daily risk-adjusted stock return by 0.0000339. In relative terms, a one-unit increase in ESG lowers the average daily risk-adjusted stock return by one percent. As presented, the coefficient is minimal; hence it can be questionable if it affects the stock return in the long run. Previous literature within the field has shown divergent results, whereas the vast majority has shown a positive relationship between ESG and financial performance, as Friede et al. (2015) stated. Our results align with Di Giuli and Kostovetsky (2014), who find a negative relation between CSR-score and stock returns.

Table 10. Results from regressions testing hypothesis II

	(1)	(2)	(3)
VARIABLES	R_RF	R_RF	R_RF
Mkt-RF	0.00961***	0.00959***	0.00961***
SMB	0.000534***	0.000468***	0.000536***
HML	0.000618***	0.000655***	0.000621***
Mom	-0.000340***	-0.000351***	-0.000337***
ESG	-3.39e-05***		
ESGF30		-0.0322***	
ESGC			-2.26e-05***
Constant	-0.00179***	-0.00376***	-0.00253***
Observations	603,840	603,840	603,840
Number of ticker	480	480	480

Notes: The results from the regressions testing hypothesis II with the different ESG variables; the ESG score, ESG factor, and the ESGC score. The stars denotes the p-value; *** p<0.01, ** p<0.05, * p<0.1

Table 11 shows the results to address hypothesis III, i.e., the four-factor alphas from the top, bottom, and difference portfolios for both 10 and 30 percent. As described in the table, all four-factor alphas are negative and significant at a one percent level. The four-factor alpha of the difference portfolio is significant at a one percent level for both portfolios. Hence, the null hypothesis can be rejected with a 99 percent confidence level, suggesting a difference in four-factor alphas between the top and bottom portfolios. Moreover, the difference portfolios are negative, and hence the bottom portfolios, both 10 and 30 percent have performed better than the top portfolios. Since both the top and bottom portfolios have negative four-factor alphas, all portfolios have underperformed compared to the market and have not generated any abnormal return when adjusted for the four factors. The results contradict our expectations as we expected that the top portfolio would perform better than the bottom, given the positive links found in previous literature (Derwall et al., 2005). Our results also differ from Eccles et al. (2014), who find that high sustainability companies outperform low sustainability companies on the stock market. Instead, our results can be explained by Wang and Sargis (2020), who stated that investors pay a slight premium for top ESG firms and hence can expect a lower return.

When the coefficient is minimal, the results could be considered a nearly no-effect scenario as described by Mănescu (2011). Mănescu (2011) claims that even if ESG is relevant for the pricing of stocks, ESG is publicly available information. If it is already incorporated in the price, there should not be any difference in risk-adjusted returns of the top and bottom portfolios. Table 11 shows a significant difference in the risk-adjusted returns for the top and bottom portfolios, which can be interpreted in different ways. If the four-factor alpha is low enough, it can be considered a no-effect scenario. Since the return variable (*R-RF*) ranges from -0.349 to 0.523, with a mean and median of -0.003 (as shown in Table 5, *Data section*), we argue that the coefficient is not small enough to be seen as a no-effect scenario.

Table 11. Four-factor alphas

Carhart four-factor		
VARIABLES	10% R-RF	30% R-RF
Top portfolio four-factor alpha	-0.00380***	-0.00379***
Bottom portfolio four-factor alpha	-0.00362***	-0.00368***
Difference four-factor alpha	-0.00410***	-0.00404***
Observations	60,384	181,152
Number of tickers	48	144

Notes: The four-factor alphas for the 10 and 30 percent portfolios testing hypothesis III. The stars denotes the p-values; *** p<0.01, ** p<0.05, * p<0.1

6.3. Robustness testing

In Table 12 and Table 13 the results from the regressions with different econometric models used to address hypothesis II and hypothesis III are presents. For both hypotheses, the robustness test indicates that the results are robust to the applied model. For hypothesis II the ESG score has a significant coefficient at a one percent level in all models. Hence, the null hypothesis can be rejected at a 99 percent confidence level, suggesting a relation between ESG score and financial performance, regardless of the applied model. Further, all the coefficients are similar for all models. The only noticeable difference is that the ESG factor is not significant in the CAPM and significant at a lower level for the Fama French three-factor model, although still negative in all models. For hypothesis III the risk-adjusted alpha is negative and significant at a one percent level regardless of the applied model. Since the difference portfolio is significant, the null hypothesis can be rejected at a 99 percent confidence level, suggesting that the risk-adjusted alphas are different regardless of the applied model.

Table 12. Robustness test, Hypothesis II

	Carhart four-factor			**Fama French three-factor**			**CAPM**		
	ESG	ESGF30	ESGC	ESG	ESGF30	ESGC	ESG	ESGF30	ESGC
VARIABLES	R_RF	R_RF	R_RF	R_RF	R_RF	R_RF	R_RF	R_RF	R_RF
Mkt-RF	0.00961***	0.00959***	0.00961***	0.00965***	0.00963***	0.00964***	0.00965***	0.00964***	0.00965***
SMB	0.000534***	0.000468***	0.000536***	0.000649***	0.000616***	0.000650***			
HML	0.000618***	0.000655***	0.000621***	0.000856***	0.000881***	0.000857***			
Mom	-0.000340***	-0.000351***	-0.000337***						
ESG	-3.39e-05***			-3.39e-05***			-3.40e-05***		
ESGF30		-0.0322***			-0.0171**			-0.00817	
ESGC			-2.26e-05***			-2.26e-05***			-2.27e-05***
Constant	-0.00179***	-0.00376***	-0.00253***	-0.00180***	-0.00376***	-0.00253***	-0.00181***	-0.00378***	-0.00255***
Observations	603,840	603,840	603,840	603,840	603,840	603,840	603,840	603,840	603,840
Number of tickers	480	480	480	480	480	480	480	480	480

Notes: The results from the regressions testing robustness for the econometric model used for hypothesis II. The stars denotes the p-values; *** p<0.01, ** p<0.05, * p<0.1

Table 13. Robustness test, Hypothesis III

	Carhart four-factor		**Fama French three-factor**		**CAPM**	
	10%	30%	10%	30%	10%	30%
VARIABLES	R-RF	R-RF	R-RF	R-RF	R-RF	R-RF
Top portfolio alpha	-0.00380***	-0.00379***	-0.00380***	-0.00380***	-0.00382***	-0.00381***
Bottom portfolio alpha	-0.00362***	-0.00368***	-0.00361***	-0.00368***	-0.00360***	-0.00369***
Difference alpha	-0.00410***	-0.00404***	-0.00411***	-0.00404***	-0.00414***	-0.00404***
Observations	60,384	181,152	60,384	181,152	60,384	181,152
Number of ticker	48	144	48	144	48	144

Notes: The results from the regressions testing robustness for the econometric model used for hypothesis III. The stars denotes the p-values; *** p<0.01, ** p<0.05, * p<0.1

6.4. General discussion

For some investors, it is important to exclude sin companies from their portfolios (Jansson & Biel, 2011), but as shown in Chart 2, screening of ESG scores is not a sufficient method to exclude sin companies; it is rather the opposite. As much as 16% of the sin companies are located in the top 10% portfolio. On the other hand, Statman and Glushkov (2009) state that investors should invest in high-scoring companies in terms of social responsibility but refrain from excluding sin companies if wanting to do good for society and well financially. Hence, it is reasonable that each investor evaluates what is important to them and does not entirely rely on the ESG score to invest in line with their beliefs.

The results from hypotheses II and III partly contradicts previous findings of a positive relation between ESG score and financial performance as summarized by Friede et al. (2015). In contrast, we find a negative relation between ESG score and financial performance. There can be many reasons for our findings. In relation to the discussion by Fatemi et al. (2018) regarding ESG disclosure for companies with ESG concerns and ESG strengths, one explanation could be that companies with high ESG scores in the sample have ESG strengths and hence disclosing is associated with more costs. Chart 3-6 shows that the top and bottom portfolios do not differ substantially in their industry composition. Hence, it is difficult to say something about the potential disadvantage of being a company with ESG strengths. Another reason could be that ESG investing is often seen as a long-term strategy, and hence the benefits might not be apparent in a time frame of five years. Further, as Mănescu (2011) stated, ESG might already be incorporated in the stock price, and hence no abnormal return exists when adjusting for risk factors.

7. Conclusion

This research contributes to a broadly studied research area regarding ESG and financial performance by focusing on the sustainable performance from ESG investing, which is a relatively unexplored field. To respond to the research question, we find evidence that the ESG score is sufficient to allocate investments to companies with top sustainable performance. This evidence goes in line with our main objective to evaluate whether ESG investing contributes to a more sustainable society.

Through this research, we can conclude that the portfolios composed of top-performing companies in terms of ESG score perform better than the portfolios composed of bottom-performing companies on most of the selected sustainability variables. Based on our findings it seems as if screening of ESG scores is a sufficient method to allocate investments to companies with top sustainable performance, and therefore serves the purpose for investors with an impact-based motive. ESG investing is, however, less efficient for investors with a financial motive, at least with a short time frame, as we find a negative, yet small, relation between ESG score and stock return during the five years studied.

Although the results in this research show that companies with high ESG scores perform better on the selected sustainability metrics, it is essential to note that it can depend on the chosen sustainability metrics, sample, and time period. We can only conclude that the ESG score can assess a company's sustainable performance for our chosen sustainability variables and chosen sample. The concern whether ESG rating is enough to evaluate a company's sustainable performance for other variables and samples remains. Hence, we agree with Schramade (2017) that investors wanting to make an impact should conduct their own research on aspects important to them. The ESG ratings are improving, and more companies report on ESG issues. However, there is a long way to go until there is a complete ESG score.

7.1. Limitations

One limitation is that ESG investing is often mentioned as a good long-term investment strategy. For example, Eccles et al. (2014) find that high sustainability companies outperform low sustainability companies in the long run. Since this thesis includes data collected for five years, there is a risk that the long-term effect has not yet been shown. Hence, it would be interesting to extend the time period.

Another limitation is related to the ESG score provider and sustainability variables. Part of the critiques against ESG scores is differences between scores depending on the rating agency. Since we only use the ESG score provided by Refinitiv, it is important to note that the results from this research are limited to one provider. Further, this research only investigates a few selected sustainability variables. The ESG score includes many variables, and if others had been chosen, the results and hence conclusion might have been different.

As pointed out by Mănescu (2011), ESG is industry-specific. More specifically, some industries might have high ESG scores and high stock returns, while other industries have low ESG scores and low stock returns. We have not specifically controlled for the industry in this study which is a limitation, instead we used firm fixed-effect.

As stated by Gregory et al. (2014), using stock market returns as the financial performance of CSR can be problematic. It is problematic since the return is expected to reflect CSR change rather than the actual score in an efficient market. Hence, the results could be misleading and underestimate corporate social performance and financial performance. It applies to this research since the ESG score is kept constant over the trading days of a calendar year, which could affect the result. However, Gregory and Whittaker (2013) point out that Corporate Social Performance is sticky, proved by a high relation between the current and a lagged CSR variable for any given year in their study.

Since a binary variable is used for several of the sustainability metrics in hypothesis I, there is not much nuance to the results. There is a risk of box-ticking with binary variables, and there can be significant differences between a 1 in one company, compared to another. Hence, there is a risk that the results are overestimated or underestimated.

7.2. Implications

The theoretical implications of this work lie in the relation between ESG score and sustainable performance. The research on sustainable performance and ESG is relatively unexplored, and hence, this paper contributes to the literature by adding another aspect to the available literature on ESG and financial performance. The practical implications of this work are based on the different motives of an investor. For investors with a financial motive, there is a need for additional research and assessment tools in order to target companies with great financial performance. On the other hand, investors with an impact-based motive can rely on ESG score screening to contribute to a more sustainable society. Our results lend support to investors that need to prove that the invested capital in their ESG portfolios makes a sustainable impact. It is, however, important to note that this research is only a starting point to assess the sustainable impact of ESG investing, and it does not give any generalized guidelines. Instead, the results are closely linked to the chosen methodology, sample, and time period. In order to make any generalized conclusions, further research is needed, as described below.

7.3. Further research

This research implies that ESG investing has a positive effect on sustainable development, but it has a negative, yet small, effect on stock return. Hence, it would be interesting to investigate the tradeoff between sustainable performance and financial performance. That is, to find the breaking point for where investors abandon ESG investing due to poor financial performance.

In addition, it is important to note that the results in this research are limited to the chosen sample, time period, and sustainability variables. Hence, it would be interesting to extend the research with another sample, time period, and sustainability variables to see if the results would diverge. Another suggestion for further research would be to investigate the effect of COVID-19 on sustainable performance to see if companies with high ESG scores handled the crisis better or worse than companies with low ESG scores. To further assess sustainable performance, it would also be interesting to look at the change in sustainable performance instead of simply looking at the values. It might show that a particular group of companies have a more remarkable change in the sustainability variables. If that were the case, it would open up for a discussion about the desired results of ESG investing; is it to assess companies that are already good in terms of sustainability or assess companies that are improving the most?

8. References

Alva Group. (2020). *What's the difference between CSR and ESG?*. Retrieved 2021-04-16 from: https://www.alva-group.com/blog/whats-the-difference-between-csr-and-esg/

Amel-Zadeh, A., & Serafeim, G. (2018). Why and how investors use ESG information: Evidence from a global survey. *Financial Analysts Journal*, *74*(3), 87-103.

Atan, R., Alam, M., Said, J., & Zamri, M. (2018). The impacts of environmental, social, and governance factors on firm performance. *Management of Environmental Quality: An International Journal, 29*(2), 182-194.

Baldini, M., Maso, L., Liberatore, D., Mazzi, G., & Terzani, F. (2018). Role of Country- and Firm-Level Determinants in Environmental, Social, and Governance Disclosure. *Journal of Business Ethics, 150*(1), 79-98.

Berg, F., Fabisik, K., & Sautner, Z. (2020a). Rewriting history II: The (un) predictable past of ESG ratings. *European Corporate Governance Institute–Finance Working Paper*, *708*.

Berg, F., Koelbel, J. F., & Rigobon, R. (2020b). *Aggregate confusion: The divergence of ESG ratings.* MIT Sloan School of Management.

Berk, J. & DeMarzo, P. (2017). *Corporate finance. 4th ed*. England: Pearson Education.

Blackwell III, J. L. (2005). Estimation and testing of fixed-effect panel-data systems. *The STATA journal*, *5*(2), 202-207.

Boffo, R., Patalano, R.. (2020). *ESG Investing: Practices, Progress and Challenges.* Retrieved from OECD Paris: https://www.oecd.org/finance/esg-investing.htm

Bolton, P., & Kacperczyk, M. (2020). *Do investors care about carbon risk?* (No. w26968). National Bureau of Economic Research.

Bos. J. (2018). *How Should Investors Measure the Societal Impact of Investing?* Retrieved from CFA Institute Magazine: https://www.cfainstitute.org/en/research/cfa-magazine/2018/how-should-investors-measure-the-societal-impact-of-investing

Brooks, C. (2019). *Introductory Econometrics for Finance 4th Edition.* Cambridge: Cambridge University Press

Brooks, C., & Oikonomou, I. (2018). The effects of environmental, social and governance disclosures and performance on firm value: A review of the literature in accounting and finance. *The British Accounting Review, 50*(1), 1-15.

Buallay, A. (2019). Is sustainability reporting (ESG) associated with performance? Evidence from the European banking sector. *Management of Environmental Quality: An International Journal*, *30*(1), 98-115.

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

Chan, L., Jegadeesh, N. & Lakonishok, J. (1999). The profitability of momentum strategies. *Financial Analysts Journal*, *55*(6), 80-90.

Cheng, B., Ioannou, I., & Serafeim, G. (2014). Corporate social responsibility and access to finance. *Strategic Management Journal*, *35*(1), 1-23.

Collins, S. & Sullivan, K. (2020). *Advancing environmental, social, and governance investing. A holistic approach for investment management firms.* Retrieved from Deloitte Insight: https://www2.deloitte.com/us/en/insights/industry/financial-services/esg-investing-performance.html

D'Aquila, J. (2018). The Current State of Sustainability Reporting: A Work in Progress. *The CPA Journal*, *88*(7), 44-50.

Derwall, J., Guenster, N., Bauer, R., & Koedijk, K. (2005). The Eco-Efficiency Premium Puzzle. *Financial Analysts Journal*, *61*(2), 51-63.

Di Giuli, A., & Kostovetsky, L. (2014). Are red or blue companies more likely to go green? Politics and corporate social responsibility. *Journal of Financial Economics*, *111*(1), 158-180.

Donthu, N., & Gustafsson, A. (2020). Effects of COVID-19 on business and research. *Journal of Business Research*, *117*, 284-289.

Drempetic, S., Klein, C., & Zwergel, B. (2019). The Influence of Firm Size on the ESG Score: Corporate Sustainability Ratings Under Review. *Journal of Business Ethics*, *167*(2), 1-28.

Eccles, G., Ioannou, I., & Serafeim, G. (2014). The impact of corporate sustainability on organizational processes and performance. *Management Science*, *60*(11), 2835-2857.

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

European Commission (2021). *Non-financial Reporting.* Retrieved 2021-04-12 from https://ec.europa.eu/info/business-economy-euro/company-reporting-and-auditing/company-reporting/non-financial-reporting_en

Fabozzi, F. J., Gupta, F., & Markowitz, H. M. (2002). The legacy of modern portfolio theory. *The Journal of Investing*, *11*(3), 7-22.

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

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

Fatemi, A., Glaum, M., & Kaiser, S. (2018). ESG performance and firm value: The moderating role of disclosure. *Global Finance Journal*, *38*, 45-64.

Fleming, R., & Ledbetter, A. (2020). Making mandatory sustainability disclosure a reality. *Environmental Law Reporter*, *50*(8), 10647-10649.

French, K (2021). Description of Fama/French Factors. Retrieved 2021-03-09 from: https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/f-f_factors.html

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

Gibson, R., Krueger, P., & Schmidt, P. S. (2019). ESG rating disagreement and stock returns. *Swiss Finance Institute Research Paper*, (19-67).

Gillan, S. L., Koch, A., & Starks, L. T. (2021). Firms and social responsibility: A review of ESG and CSR research in corporate finance. *Journal of Corporate Finance, 66 (101889)*

Gregory, A., Tharyan, R., & 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.

Gregory, A., & Whittaker, J. (2013). Exploring the valuation of corporate social responsibility--A comparison of research methods. *Journal of Business Ethics, 116*(1), 1-20.

Jansson, M., & Biel, A. (2011). Motives to engage in sustainable investment: A comparison between institutional and private investors. *Sustainable Development, 19*(2), 135-142.

Jegadeesh, N., & Titman, S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *Journal of Finance, 48*(1), 65-91.

Jensen, J., & Berg, N. (2012). Determinants of Traditional Sustainability Reporting Versus Integrated Reporting. An Institutionalist Approach. *Business Strategy and the Environment, 21*(5), 299-316.

Li, Y., Gong, M., Zhang, X. Y., & Koh, L. (2018). The impact of environmental, social, and governance disclosure on firm value: The role of CEO power. *The British Accounting Review*, *50*(1), 60-75.

Liang, H., & Renneboog, L. (2017). On the foundations of corporate social responsibility. *The Journal of Finance*, *72*(2), 853-910.

Lintner, J. (1965). The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *The Review of Economics and Statistics, 47*(1), 13-37.

Malkiel, B. G., & Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, *25*(2), 383-417.

Mănescu, C. (2011). Stock returns in relation to environmental, social and governance performance: Mispricing or compensation for risk?. *Sustainable Development*, *19*(2), 95-118.

Markowitz, H., (1952). Portfolio Selection. *The Journal of Finance,* 7(1), pp.77-91

Matos, P. (2020). *ESG and Responsible Institutional Investing Around the World: A Critical Review.* CFA Institute Research foundation.

Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica (pre-1986), 34*(4), 768-783.

MSCI. (2021). *GICS - The Global Industry Classification Standard.* Retrieved 2021-04-08 from: https://www.msci.com/gics

OECD (2020), *OECD Business and Finance Outlook 2020: Sustainable and Resilient Finance.* OECD Publishing, Paris. Retrieved from: https://doi.org/10.1787/eb61fd29-en.

Perold, A. (2004). The Capital Asset Pricing Model. *Journal of Economic Perspectives, 18*(3), 3-24.

Plumlee, M., Brown, D., Hayes, R., & Marshall, R. (2015). Voluntary environmental disclosure quality and firm value: Further evidence. *Journal of Accounting and Public Policy, 34*(4), 336-361.

PwC. (2020). *2022 - The growth opportunity of the century. Are you ready for the ESG change?* Retrieved from PwC Luxembourg: https://www.pwc.lu/en/sustainable-finance/esg-report-the-growth-opportunity-of-the-century.html

Refinitiv. (2021). *Environmental, Social and Governance (ESG) Scores from Refinitiv.* Retrieved from Refinitiv: https://www.refinitiv.com/content/dam/marketing/en_us/documents/methodology/refinitiv-esg-scores-methodology.pdf

Runesson, E., Samani, N. & Marton, J. (2018). *Financial Accounting Theory - An Accounting Quality Approach*. Lund: Studentlitteratur AB.

Rupley, K., Brown, D., & Marshall, S. (2017). Evolution of corporate reporting: From stand-alone corporate social responsibility reporting to integrated reporting. *Research in Accounting Regulation*, *29*(2), 172-176.

Sahut, J., & Pasquini-Descomps, H. (2015). ESG Impact on Market Performance of Firms: International Evidence. *Management International*, *19*(2), 40-63.

Salkind, N. J. (2010). *Encyclopedia of research design* (Vols. 1-0). Thousand Oaks, CA: SAGE Publications, Inc. doi: 10.4135/9781412961288

Schramade, W. (2017). Investing in the UN sustainable development goals: opportunities for companies and investors. *Journal of Applied Corporate Finance*, *29*(2), 87-99.

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

Shiller, R. J. (2013). *Finance and the Good Society*. Princeton: Princeton University Press.

Statman, M., & Glushkov, D. (2009). The wages of social responsibility. *Financial Analysts Journal*, *65*(4), 33-46.

Treynor, J. L. 1962. "Toward a Theory of Market Value of Risky Assets." Unpublished manuscript. Final version in *Asset Pricing and Portfolio Performance*, 1999, Robert A. Korajczyk, ed., London: Risk Books, pp. 15–22

United Nations (2021). *The 17 goals, History.* Retrieved 2021-03-17 from https://sdgs.un.org/goals

Vanwalleghem, D., & Mirowska, A. (2020). The investor that could and would: The effect of proactive personality on sustainable investment choice. *Journal of Behavioral and Experimental Finance*, 26(100313)

Velte, P. (2017). Does ESG performance have an impact on financial performance? Evidence from Germany. *Journal of Global Responsibility*, *8*(2), 169-178.

Verheyden, T., Eccles, R. G., & Feiner, A. (2016). ESG for all? The impact of ESG screening on return, risk, and diversification. *Journal of Applied Corporate Finance*, *28*(2), 47-55.

Wang, P. & Sargis, M. (2020). *Better Minus Worse - Evaluating ESG Effects on Risk and Return.* (Morningstar Quantitative Research, 1). Retrieved from: https://www.morningstar.com/insights/2020/02/19/esg-companies

9. List of tables

10. List of charts

11. List of equations

12. Appendix

Appendix A. Efficient market hypothesis

The efficient market hypothesis, EMH, is a theory that was introduced by Malkiel and Fama in 1970 and has since then been widely used in financial literature. The theory states that all available information should be included in the stock price, reflecting the fair value of the stock (Malkiel & Fama, 1970). If the hypothesis holds, the market is said to be efficient, and it would be impossible to beat the market without taking on higher risk.

Appendix B. Modern portfolio theory

According to the EMH, it is only possible to beat the market if the investor takes on higher risk. With higher risk comes higher expected returns (Malkiel & Fama, 1970). This relation is discussed in an article on portfolio selection, published by Harry Markowitz, which laid the foundation for the Modern Portfolio Theory in 1952 (Fabozzi, Gupta & Markowitz, 2002). The theory is about finding the optimal portfolio composition given a certain level of risk. That is, to find the optimal trade-off between return and risk in a portfolio. The theory states that a rational investor will always choose a portfolio with lower risk over a portfolio with higher risk, given the same expected return. The theory further states that the investor can reduce the risk in the portfolio through diversification (Markowitz, 1952).

Appendix C. Carhart four-factor model

The Carhart four-factor model is based on the Fama French three-factor model but with an additional factor that captures a one-year momentum abnormality, developed by Jegadesh and Titman (1993). A momentum trading strategy is about buying past winners and selling past losers, and Jegadeesh and Titman (1993) find significant abnormal returns during their studied period. The momentum factor was added by Carhart (1997), given the inability in the three-factor model to explain cross-sectional variation in momentum-sorted portfolio returns (Carhart, 1997). The momentum abnormality is explained as market inefficiency due to the slow reaction to incorporate the full impact of information into valuations (Chan, Jegadeesh & Lakonishok, 1999); however, it is robust to time periods (Jegadeesh & Titman, 1993).

$$r_{i,t} - rf_t = \alpha_{i,t} + \beta_{1,i}(r_{mkt,t} - rf_t) + \beta_{2,i}SMB_t + \beta_{3,i}HML_t + \beta_{4,i}MOM_t + \varepsilon_{i,t}$$

Equation 7

$r_{i,t}$ = *return of the stock i at time t*
i = the i:th stock
t = time t
$r_{i,t}$ = *return of the stock i at time t*
rf_t = *risk-free rate at time t*
$\alpha_{i,t}$ = *alpha, the risk-adjusted abnormal return for stock i*
$\beta_{1-4,i}$ = *the betas of the regressors*
$(r_{mkt,t} - rf_t)$ = *the difference in expected return of the market minus the risk-free rate in time t, the risk premium*
SMB_t = *The size premium at time t*
HML_t = *The value premium at time t*
MOM_t = *The momentum factor at time t*
$\varepsilon_{i,t}$ = *error term for stock i at time t*

Source: Brooks (2019)

Appendix D. Definition of the chosen sustainbility variables

ENVIRONMENTAL	SOCIAL	GOVERNANCE
CO2 Equivalent Emissions/ Million in Revenue $ (Emissions)	**Policy responsible marketing (Product responsibility)**	**Policy board independence (Management)**
Total CO2 and CO2 equivalents emission in tonnes divided by net sales or revenue in US dollars. Total CO2 emission = direct (scope 1) + indirect (scope 2) emissions. **NOTE:** For total CO2 and CO2 equivalents emission in tonnes, the following gases are relevant: carbon dioxide (CO2), methane (CH4), nitrous oxide (N2O), hydrofluorocarbons (HFCS), perfluorinated compound (PFCS), sulfur hexafluoride (SF6), nitrogen trifluoride (NF3). Refinitiv follows the greenhouse gas (GHG) protocol for all their emission classifications by type.	Define if the company has a policy on responsible marketing, ensuring the protection of children. This includes information on responsible marketing communications approaches undertaken by companies in order to protect weak customers like children. Legal obligations information is not considered.	Define if the company has a policy regarding the independence of its board. If yes, the company strives to maintain a well-balanced board through an adequate number of independent board members. The independent board members maintain integrity and independence in decision making.
Environmental products (Environmental innovation score)	**Policy customer health and safety (Product responsibility)**	**Shareholder rights policy (Shareholder score)**

Defines if the company reports on at least one product line or service designed to have positive effects on the environment or which is environmentally labeled and marketed. The products and services that have positive environmental effects or marketed as something that solve environmental problems are in focus.	Defines if the company has a policy to protect customer health & safety. More specifically, if they have processes or initiatives in place by which they strive to market products that are fostering benefits to the consumer's health & safety rather than putting it at risk.	Defines if the company has a policy for ensuring equal treatment of minority shareholders, facilitating shareholder engagement, or limiting the use of anti-takeover devices.
Env supply chain partnership termination (Resource use)	**Net employment creation (Workforce score)**	**Board gender diversity (Management)**
Defines if the company reports or shows that they are ready to end a partnership with a sourcing partner if environmental criteria are not met.	Total employment growth over the last year in percentage.	Percentage of females on the board.
Environmental partnership (Emission)		
Define if the company reports on partnerships or initiatives with specialized NGOs, industry organizations, governmental or supra-governmental organizations focusing on improving environmental issues.		

Notes: The definitions from the Refinitiv Eikon database of the 10 sustainability variables used in the thesis. The variables are divided into each of the ESG pillars, and the category that the variable belongs to is presented in the parenthesis.

Appendix E. Equations for the factors from Kenneth R. French data library

The SMB, Small Minus Big;

$$SMB = 1/3(Small\ Value + Small\ Neutral + Small\ Growth) - 1/3(Big\ Value + Big\ Neutral + Big\ Growth)$$

Equation 8

The HML, High minus Low;

$$HML = 1/2(Small\ Value + Big\ Value) - 1/2(Small\ Growth + Big\ Growth)$$

Equation 9

The Mom, daily Momentum factor;

$$Mom = 1/2(Small\ High + Big\ High) - 1/2(Small\ Low + Big\ Low)$$

Equation 10

Appendix F. Yearly industry distribution of the portfolios

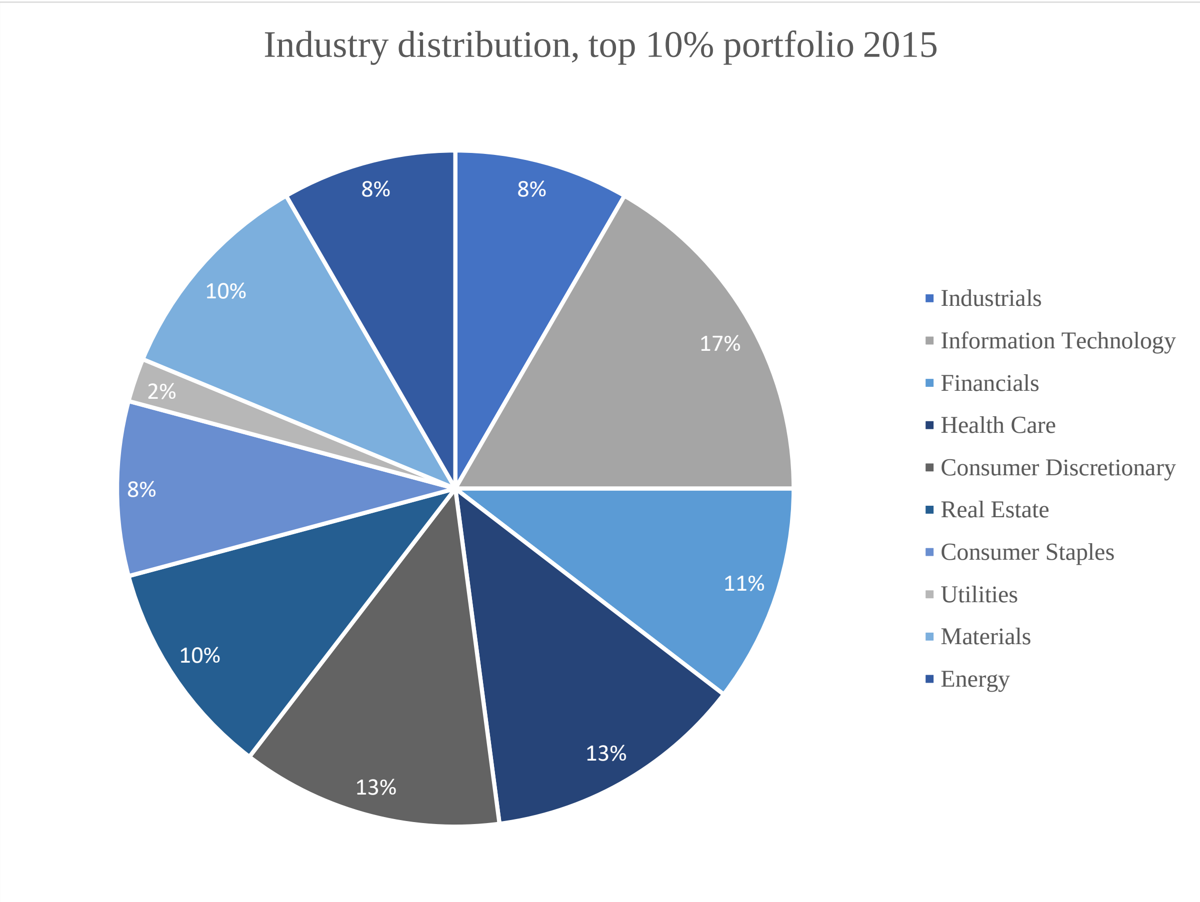
Table 14. Industry distribution of the top 10% portfolio

Top 10 % portfolio

GICS Sector name	**2015**	**2016**	**2017**	**2018**	**2019**
Industrials	8.3%	12.5%	14.6%	8.3%	6.3%
Information Technology	16.7%	10.4%	8.3%	12.5%	12.5%
Financials	10.4%	12.5%	12.5%	8.3%	10.4%
Health Care	12.5%	16.7%	14.6%	10.4%	10.4%
Consumer Discretionary	12.5%	18.8%	16.7%	20.8%	16.7%
Real Estate	10.4%	8.3%	8.3%	10.4%	12.5%
Consumer Staples	8.3%	10.4%	10.4%	12.5%	14.6%
Utilities	2.1%	4.2%	6.3%	0.0%	0.0%
Communication Services	0.0%	0.0%	0.0%	0.0%	0.0%
Materials	10.4%	6.3%	6.3%	10.4%	8.3%
Energy	8.3%	0.0%	2.1%	6.3%	8.3%
Total	100%	100%	100%	100%	100%

Notes: The company distribution in the 11 GICS sectors for the top 10 percent portfolio for each of the studied years.

Chart 3. Industry distribution of the top 10% portfolio 2015



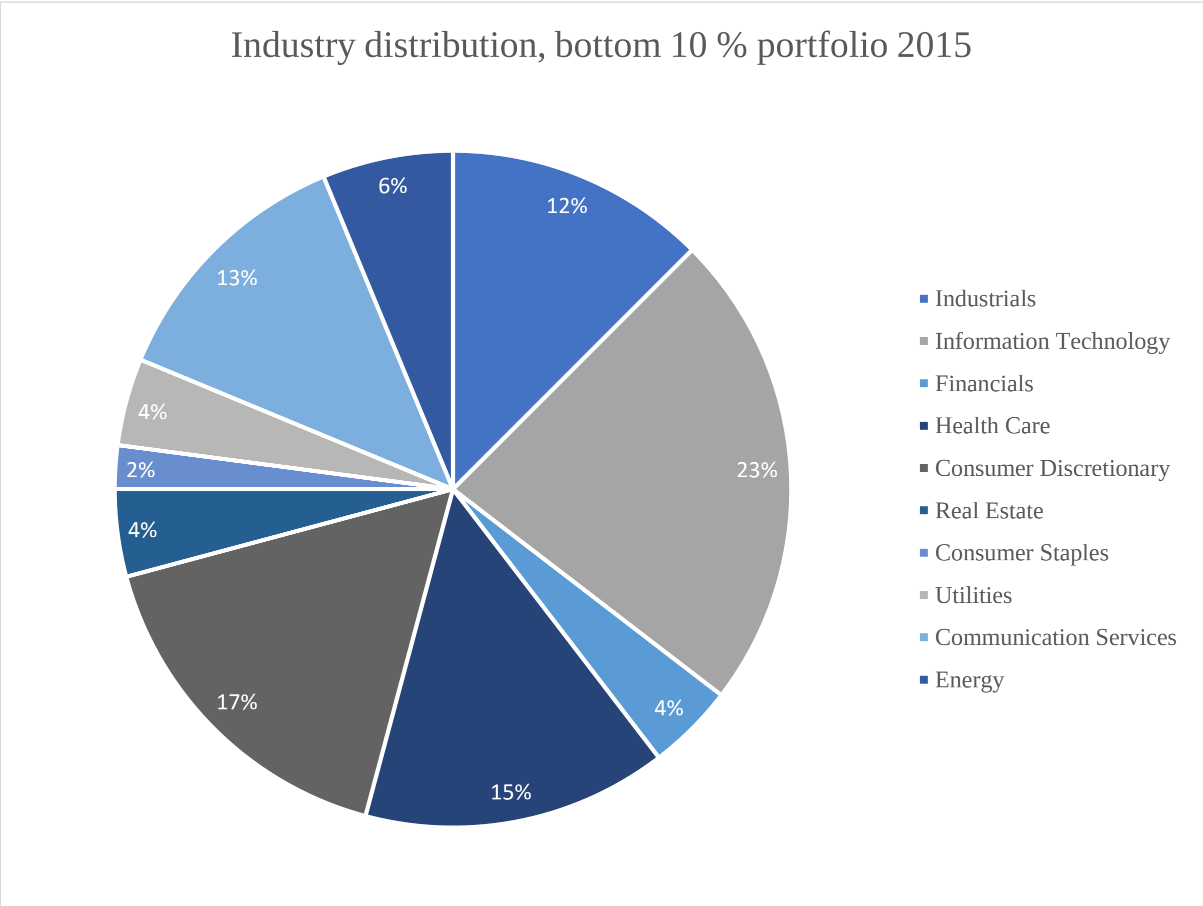
Notes: The top 10 percent portfolio (48 companies) divided in the 11 GICS sectors, year 2015.

Table 15. Industry distribution of the bottom 10% portfolio

Bottom 10 % portfolio					
GICS Sector name	**2015**	**2016**	**2017**	**2018**	**2019**
Industrials	12.5%	14.6%	14.6%	16.7%	14.6%
Information Technology	22.9%	25.0%	22.9%	29.2%	20.8%
Financials	4.2%	2.1%	2.1%	6.3%	4.2%
Health Care	14.6%	12.5%	12.5%	8.3%	14.6%
Consumer Discretionary	16.7%	16.7%	16.7%	18.8%	18.8%
Real Estate	4.2%	6.3%	4.2%	2.1%	4.2%
Consumer Staples	2.1%	2.1%	4.2%	2.1%	4.2%
Utilities	4.2%	4.2%	4.2%	2.1%	0.0%
Communication Services	12.5%	12.5%	14.6%	10.4%	14.6%
Materials	0.0%	0.0%	0.0%	0.0%	0.0%
Energy	6.3%	4.2%	4.2%	4.2%	4.2%
Total	100%	100%	100%	100%	100%

Notes: The company distribution in the 11 GICS sectors for the bottom 10 percent portfolio for each of the studied years.

Chart 4. Industry distribution of the bottom 10% portfolio 2015



Notes: The bottom 10 percent portfolio (48 companies) divided in the 11 GICS sectors, year 2015.

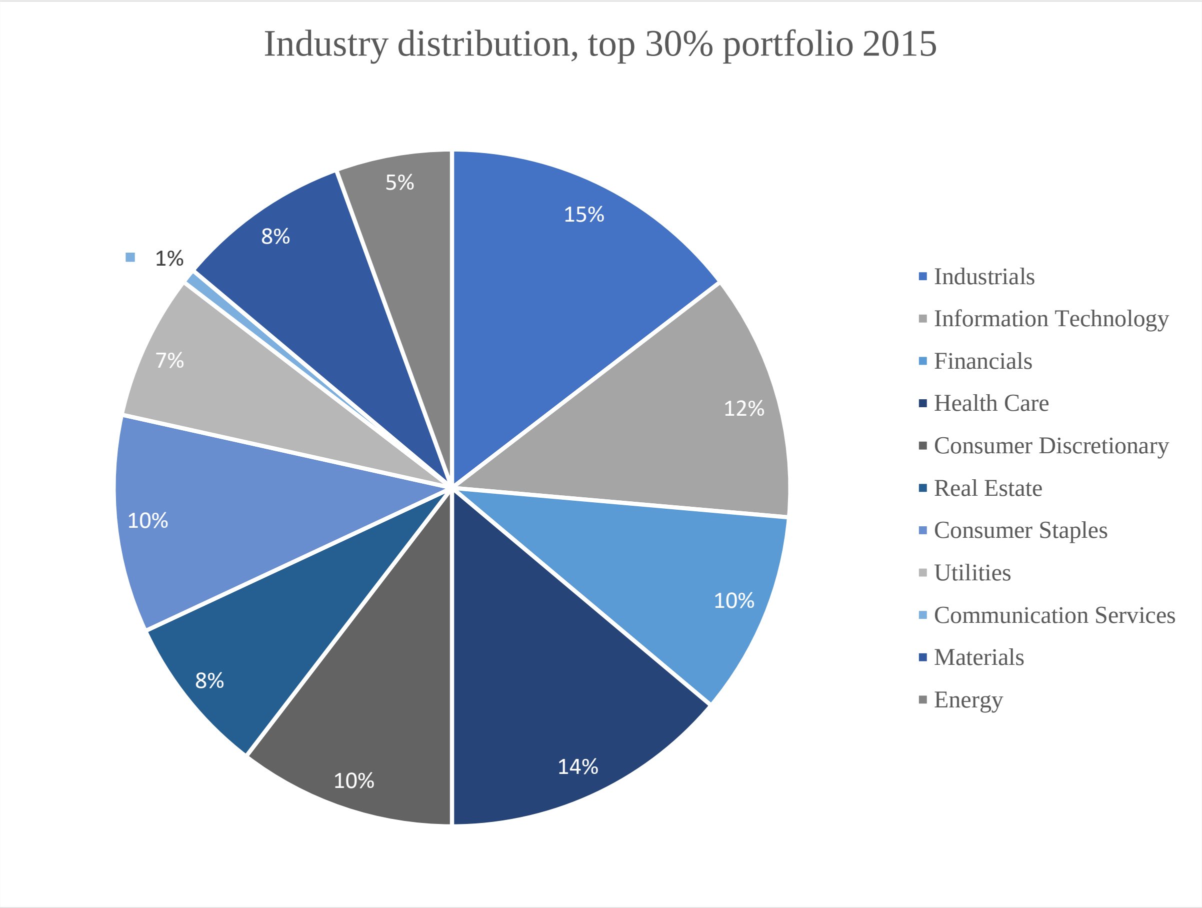
Table 16. Industry distribution of the top 30% portfolio

Top 30 % portfolio

GICS Sector name	**2015**	**2016**	**2017**	**2018**	**2019**
Industrials	14.6%	12.5%	13.9%	10.4%	10.4%
Information Technology	11.8%	13.9%	14.6%	14.6%	13.2%
Financials	9.7%	9.0%	9.0%	11.8%	10.4%
Health Care	13.9%	15.3%	13.9%	14.6%	15.3%
Consumer Discretionary	10.4%	9.7%	11.8%	12.5%	10.4%
Real Estate	7.6%	6.9%	9.7%	9.0%	7.6%
Consumer Staples	10.4%	12.5%	9.7%	9.7%	11.8%
Utilities	6.9%	7.6%	5.6%	4.9%	6.3%
Communication Services	0.7%	0.0%	0.7%	0.7%	1.4%
Materials	8.3%	6.9%	6.3%	6.3%	6.9%
Energy	5.6%	5.6%	4.9%	5.6%	6.3%
Total	100%	100%	100%	100%	100%

Notes: The company distribution in the 11 GICS sectors for the top 30 percent portfolio for each of the studied years.

Chart 5. Industry distribution of the top 30% portfolio 2015



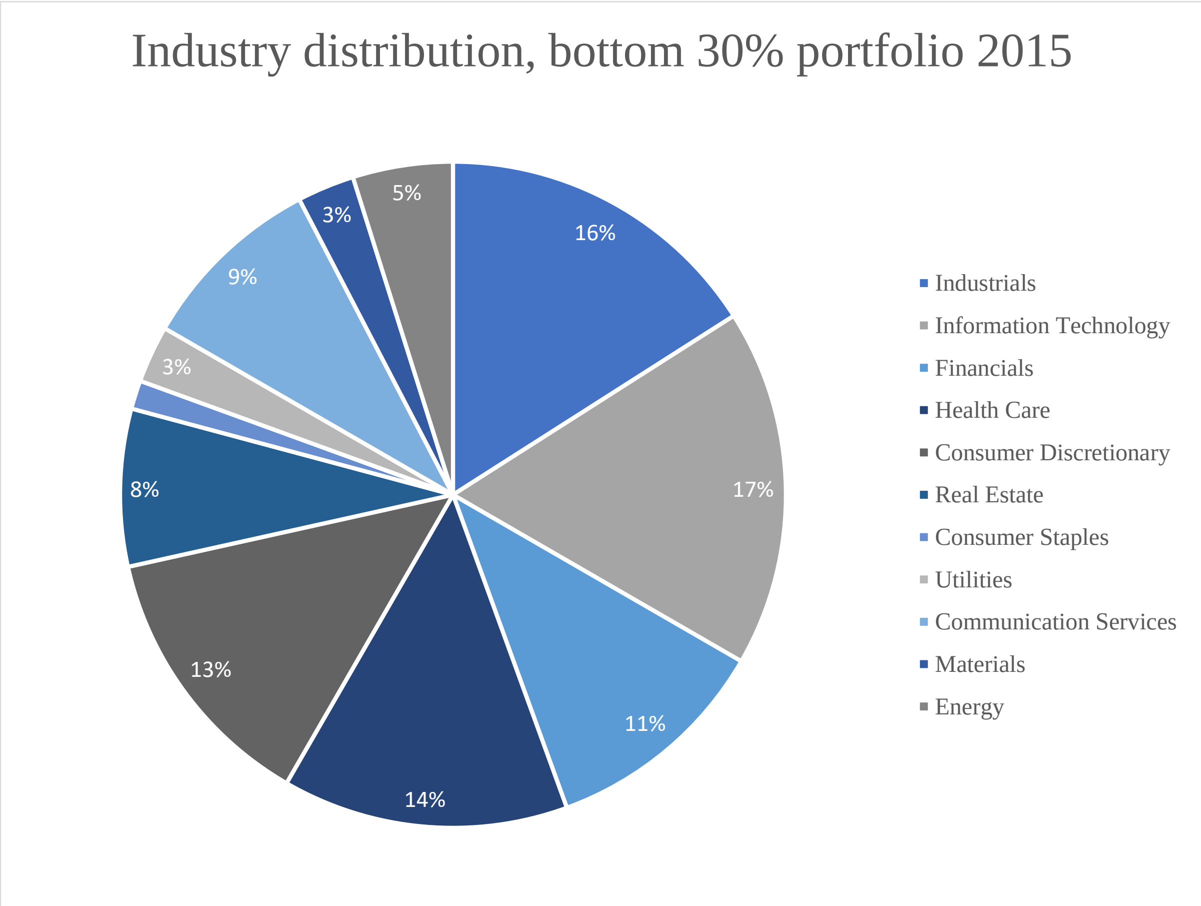
Notes: The top 30 percent portfolio (144 companies) divided in the 11 GICS sectors, year 2015.

Table 17. Industry distribution of the bottom 30% portfolio

Bottom 30 % portfolio					
GICS Sector name	**2015**	**2016**	**2017**	**2018**	**2019**
Industrials	16.0%	16.0%	16.0%	13.9%	14.6%
Information Technology	17.4%	17.4%	18.1%	17.4%	18.8%
Financials	11.1%	11.1%	13.2%	15.3%	16.7%
Health Care	13.9%	14.6%	12.5%	11.8%	8.3%
Consumer Discretionary	13.2%	13.2%	15.3%	17.4%	16.7%
Real Estate	7.6%	6.9%	5.6%	4.2%	2.1%
Consumer Staples	1.4%	1.4%	1.4%	1.4%	1.4%
Utilities	2.8%	3.5%	3.5%	2.8%	3.5%
Communication Services	9.0%	9.0%	7.6%	8.3%	9.0%
Materials	2.8%	2.1%	2.8%	3.5%	4.9%
Energy	4.9%	4.9%	4.2%	4.2%	4.2%
Total	100%	100%	100%	100%	100%

Notes: The company distribution in the 11 GICS sectors for the bottom 30 percent portfolio for each of the studied years.

Chart 6. Industry distribution of the bottom 30% portfolio 2015



Notes: The bottom 30 percent portfolio (144 companies) divided in the 11 GICS sectors, year 2015.

Appendix G. Capital Asset Pricing Model, CAPM

The capital asset pricing model, CAPM, was the first coherent framework that aimed to answer how the risk of an investment should affect the expected return of the investment (Perold, 2004). Different authors developed the model in the 1960s; Sharpe (1964), Treynor (1962), Lintner (1965), and Mossin (1966). The model revolves around the idea that not all risks should affect the pricing of an asset. More specifically, a risk that can be diversified away in a portfolio should not affect the pricing since it is not a real risk (Perold, 2004). Like all models, it is not perfect and relies on certain assumptions; however, the CAPM is the most common and important risk and return model (Berk & DeMarzo, 2017).

The model states that the expected return of an investment is the rate of a risk-free security and a risk premium.

$$r_{i,t} - rf_t = \alpha_{i,t} + \beta_{1,i}(r_{mkt,t} - rf_t) + \varepsilon_{i,t}$$

Equation 11

$r_{i,t}$ = *return of the stock i at time t*
$i = the\ i{:}th\ stock$
$t = time\ t$
$r_{i,t} = return\ of\ the\ stock\ i\ at\ time\ t$
$rf_t = risk - free\ rate\ at\ time\ t$
$\alpha_{i,t} = alpha, the\ risk - adjusted\ abnormal\ return\ for\ stock\ i$
$\beta_{1,i} = the\ beta\ of\ the\ regressor$
$(r_{mkt,t} - rf_t)$ = *the difference in expected return of the market minus the risk-free rate in time t, the risk premium.*
$\varepsilon_{i,t} = error\ term\ for\ stock\ i\ at\ time\ t$

Source: Brooks (2019)

Appendix H. Fama French three-factor model

The Fama French three-factor model is a development of the CAPM, where the authors have added size risk (Market equity, ME) and value risk factors (book-to-market equity, BE/ME) to the market risk factor (Fama & French, 1992). The size risk, ME, is the market capitalization, meaning a stock's price times the number of shares outstanding. The value risk factor, BE/ME, is calculated as the book value of equity divided by the market value of equity. This factor is also known as value versus growth stocks. There is a lag in matching the accounting (book) values with the market values of a minimum of six months to ensure that the accounting variables are known before the returns they are used to explain.

Through the variable SMB, Small minus Big, the authors aim to mimic the size risk related to return. It is calculated as the difference between the returns on three small-stock and three big-stock portfolios with similar weighted average book-to-market equity (Fama & French, 1993). Instead, the variable HML, High minus Low, aims to mimic the risk factor related to book-to-market equity. It is calculated as the difference between the simple average of portfolios, two portfolios including companies with high BE/ME and two with low BE/ME with similar weighted-average size (ibid.).

$$r_{i,t} - rf_t = \alpha_{i,t} + \beta_{1,i}(r_{mkt,t} - rf_t) + \beta_{2,i}SMB_t + \beta_{3,i}HML_t + \varepsilon_{i,t}$$

Equation 12

$r_{i,t}$ = *Return of the stock i at time t*
i = the i:th stock
t = time t
$r_{i,t}$ = *return of the stock i at time t*
rf_t = *risk-free rate at time t*
$\alpha_{i,t}$ = *alpha, the risk-adjusted abnormal return for stock i*
$\beta_{1-3,i}$ = *the betas of the regressors*
$(r_{mkt,t} - rf_t)$ = *the difference in expected return of the market minus the risk-free rate in time t, the risk premium*
SMB_t = *The size premium at time t*
HML_t = *The value premium at time t*
$\varepsilon_{i,t}$ = *error term for stock i at time t*

Source: Brooks (2019)

Appendix I. OLS assumptions

Technical notation	*Interpretation*
(1) $E(u_t) = 0$	The errors have zero mean
(2) $var(u_t) = \sigma^2 < \infty$	The variance of the errors is constant and finite over all values of x_t
(3) $cov(u_i, u_j) = 0$	The errors are linearly independent of one another
(4) $cov(u_t, x_t) = 0$	There is no relationship between the error and corresponding x variate
(5) $u_t \sim N(0, \sigma^2)$	- i.e., that u_t is normally distributed

Notes: The OLS assumptions technical notation and interpretation of the assumptions.