UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

Bachelor's project within the Business and Economics Bachelor Programme

Beyond Profits: Exploring the Investment Styles and Risk-Adjusted Returns of ESG-Driven Portfolios

Abstract:
This study uses daily data to examine how different ESG implementations affect performance and portfolio characteristics. With a non-homogenous view of how ESG investing is defined, ten different value-weighted portfolios are constructed. The geographical focus is the US market, with the S&P 500 total return index (SPXTR) as the screening universe. We use different methods to assess the research questions, with the Carhart four-factor model being the main one. To account for eventual time-varying aspects, the sample is split into two periods, 2007-2014 and 2015-2022. The results suggest that pursuing a non-ESG approach is overall better for risk-adjusted returns. Furthermore, some investors may wish to implement ESG into their portfolios. If this is the case, a negative screening technique is more advantageous for risk-adjusted returns. The results also indicate that pursuing specific ESG approaches will imply different portfolio tilts. Moreover, the results suggest that in general, the ESG portfolios are tilted toward larger capitalization and value stocks, while non-ESG portfolios are tilted toward growth stocks and smaller firms. Somewhat time-varying results are found. However, the results largely remain the same when the time period is split.

Bachelor thesis 15 hp
Authors: Jonathan Taimory
Alexander Olsson
Supervisor: Charles Nadeau
Spring term 2023

JEL classifications: A22, C01, C87, G11, G12, O51, Q01

Keywords: Environmental, Social and Governance (ESG), Risk-adjusted Returns, Investment Styles

Acknowledgements

We want to express our sincerest gratitude to our supervisor Charles Nadeau for the invaluable guidance and support during the thesis. Without your expertise, patience and commitment, this thesis would not have been possible. Furthermore, we want to thank Aineas Mallios for all the exceptional assistance with the data and method section.

Table of Contents

1. Introduction

This section begins by providing a background to familiarize the reader with the overall subject matter of the thesis. Subsequently, the problem description establishes a connection between the background and the thesis. In parallel, the aim, together with the research questions, are stated as well as the contributions to the already existing literature. Finally, the section concludes with an outline of the remaining parts of the thesis.

1.1 Background Description

Investing using environmental, social and governance criteria has been a theme with growing interest. Research trying to examine the ESG criteria and financial performance started already at the beginning of the 1970s (Friede, Busch and Bassen 2015). Since the 1970s, numerous studies have been done. However, even if the literature is comprehensive, it is still not unambiguous (Henriksson et al. 2019). Between 2014 and 2018 approximately five billion dollars per year was flowing into US sustainable funds. In 2019 that number increased to 20 billion dollars and in 2020, it more than doubled to 50 billion dollars. There is a clear trend of capital flowing into firms that fulfill specific ESG criteria. Preferences for investing in firms with high ESG-rating have been communicated by institutional investors both implicitly and explicitly (Lioui and Tarelli 2022). Wu (2022) express in JP Morgan's ESG outlook of January 2022 that sustainable investing is here to stay and will continue to grow. It is stated how a combination of factors will support the growth of sustainable investing. These factors are, for example, increasing demand from investors and improvement in data provision. The increased importance of ESG investing displays the relevance of understanding the implications for investors.

Before proceeding, we want to present a note to the readers. ESG investing is a term that includes a wide range of nomenclature. This mixed terminology can sometimes lead to confusion (Grim and Berkowitz 2020). To make it more convenient, a table of various terms used in this thesis is exhibited in Appendix A, table A1.

1.2 Problem Description and Problem Analysis

As mentioned above, implementing ESG as a part of the investment strategy has been a strong trend for the last decade. There has been much research regarding ESG, which is a natural response to the increased importance of the topic.

A considerable amount of the research has focused on the trade-off between incorporating ESG into the investment strategy versus omitting it (see: Ibikunle and Steffen 2015). However, the conclusions differ and often depend on aspects such as assumptions, time periods, and methods used. Furthermore, many papers examine this relationship by constructing one ESG portfolio and one non-ESG and/or conventional portfolio (see: Derwall et al. 2005). This implicitly leads to treating ESG investing as a phenomenon that is implemented in a homogenous way. We would argue that incorporating ESG into the investment strategy can have many dimensions. For example, it is possible to implement a strategy that invests in all industries but solely the best in class firms (see: Dorfleitner, Kreuzer and Sparrer 2020). Another alternative can be divesting all firms in non-ESG industries (see: Barnett and Salomon 2006). There is still an ongoing debate about how these different implementations of ESG investing will affect portfolios.

1.3 Aim of the Study and Contributions

Not only has the prior research focused a lot on an eventual trade-off (Friede, Busch and Bassen 2015). Furthermore, the emphasis has often been on a few strategies. For example, Nagy, Kassam and Lee (2016) examine the effect on risk-adjusted returns from two strategies. The first strategy is ESG tilt which means overweighting companies with higher ESG ratings. ESG momentum is the second strategy. The strategy builds upon overweighting those companies that have recently enhanced their ESG ratings. This creates opportunities to investigate and compare more approaches, such as, for example, portfolios building upon each individual pillar of ESG. The purpose of this thesis is, therefore, to examine how different ESG approaches affect portfolio performance in the US market. All the specific ESG approaches will be presented in detail in sections 3 and 4.

Formally the first research question is:

I: How do the risk-adjusted returns differ between the constructed portfolios?

Pursuing different implementations of ESG investing can lead to particular portfolio tilts. In the financial literature, these tilts are often referred to as investment styles. According to Kumar (2009) investors can have different preferences regarding portfolio styles. For example, some investors prefer so-called value stocks, while others prefer growth stocks.

As demonstrated above, much research focuses on if it is possible to pursue ESG investing without hurting performance. However, different types of ESG investing may implicitly lead to particular portfolio tilts. As far as we are aware, this question tends to receive less focus. Nonetheless, it can be of value for investors to understand what kind of tilts different ESG approaches can lead to. Our desire is that the results can be used as a potential framework for investors considering how to incorporate ESG into their investment strategy.

Formally the second research question is:

II: Is there any difference in investment style between the portfolios? If so, how does the investment style differ?

Our thesis contributes to the literature by adopting a non-homogenous perspective on ESG investing and employing ten different portfolios. By analyzing more ESG approaches in the US market, we hope to increase the understanding of what potential implications different strategies have. Furthermore, most of the studies use data on a monthly basis (see: Auer 2014; Dorfleitner, Kreuzer and Sparrer 2020). By using daily data, we desire to increase the statistical significance of the results, making them useful for further analysis.

1.4 Outline of the Thesis

The remainder of the thesis is structured as follows. Section 2 presents the theoretical framework together with a review of previous research. Section 3 outlines the data-generating process together with the estimation model. Section 4 describes the methodology used to examine the research questions. Section 5 presents the results and discusses the main findings. Section 6 outlines the conclusions together with suggestions for future research.

2. Theoretical Framework and Literature Review

This section provides an overview of the theoretical framework later used in the thesis. It also encompasses a review of prior research conducted in the field. The aim is to establish a link between the theoretical framework, prior findings and the research questions of this thesis.

2.1 Environmental, Social and Corporate Governance (ESG)

ESG is a shortening for Environmental, Social and Corporate Governance. Each of the three pillars delves into one distinct aspect of sustainability. The environmental pillar is connected to the firm's effectiveness in managing its operations concerning emissions, pollution and depletion of natural resources. The social pillar scrutinizes the firm´s approach to its business relationships. This can involve its commitment to topics such as worker safety, human rights issues, and involvement in social communities. Lastly, the governance pillar assesses whether the firm exhibits sustainable management practices that shield shareholder rights. It is also tied to the company's internal control systems that are put in place in order to ensure compliance with laws and regulations governing its operations (Bender et al. 2017).

The incorporation of the ESG concept in investment decisions among today's investors is most commonly displayed in the form of ESG scores. Various private institutes, such as MSCI, Morningstar, Sustainalytics and Refinitiv Eikon, analyze and assign ESG scores to companies. These are based on how firms perform concerning the three sustainability criteria. Furthermore, as shown in the literature review, there are a substantial number of various methodologies for calculating an adequate ESG score. Differences arise in what the three factors consist of and how they are weighted. However, the final score is always a composition of the three criteria. Nevertheless, these differences contribute to one of the considerable criticisms of ESG scores, namely the lack of a uniform framework when grading firms. The consequence is that the correlation between ESG scores provided by the different institutes is relatively low (Boffo and Patalano 2020).

2.2 Modern Portfolio Theory (MPT)

Modern Portfolio Theory provides a framework for constructing portfolios by considering the relationship between risk and return. At its core, it emphasizes the importance of diversification to manage risk effectively.

By combining assets with different risk and return characteristics, investors can reduce the overall risk of the entire portfolio without sacrificing potential returns. Furthermore, the theory highlights the trade-off between risk and return, recognizing that investors should be compensated with higher returns for taking on higher levels of risk. One important assumption that is made is that investors are risk-averse. Thus, every rational investor will choose the portfolio with the lowest risk given a level of expected return.

While it has its limitations and relies on certain assumptions, Modern Portfolio Theory has influenced portfolio management practices and provided a foundation for understanding the risk-return relationship in investment decision-making (Markowitz 1952).

2.3 Capital Asset Pricing Model (CAPM)

The capital asset pricing model is one of the pillars of financial economics and was developed by Sharpe (1964), Lintner (1965) and Mossin (1966), building upon the ideas of Markowitz (1952). The model describes the relationship between systematic risk, also called non-diversifiable risk, and expected return. In CAPM, there is only one factor that determines the expected return of an asset and that is the systematic risk. Beta measures the systematic risk for an individual security or portfolio. Firm-specific risk is not priced by the market and thus, an investor does not get compensation for this risk. The market portfolio has a beta of one. Usually, betas that are greater than one are considered aggressive stocks. In contrast, betas that are smaller than one are considered defensive (Bodie, Kane and Marcus. 2020).

The formal equation for beta is:

$$\beta_i = \frac{Cov(R_i, R_m)}{\sigma_m^2} \quad (1)$$

Where:

$Cov(R_i, R_m) = Covariance\ between\ the\ return\ of\ asset\ i\ and\ the\ market\ portfolio$

$\sigma_m^2 = Variance\ of\ the\ market\ portfolio$

The formal equation for CAPM is:

$$E\left[R_i\right] = R_f + \beta_i(E\left[R_m - R_f\right]) \quad (2)$$

Where:

$E\left[R_i\right] = Expected\ return\ of\ asset\ i$

$R_f = Risk - free\ rate$

$\beta_i = Beta\ of\ asset\ i$

$E\left[R_m - R_f\right] = Expected\ excess\ return\ of\ the\ market\ portfolio$

Numerous anomalies have been discovered in empirical research that disagrees with CAPM. One of these examples is Banz (1981) which reveals the impact of company size. The findings suggest that smaller-sized companies exhibit abnormal returns and that CAPM is not able to measure the actual risk. It is therefore suggested that the risk of stocks is multidimensional rather than one-dimensional (Fama and French 1992). There are several alternatives to CAPM that try to account for the anomalies discovered in the research. Some of the most common ones are the Fama-French three-factor model (Fama and French 1992), the Carhart Four-factor model (Carhart 1997) and the Fama-French five-factor model (Fama and French 2015).

2.4 Literature Review

As previously discussed, the literature on ESG and financial performance goes back to the early 1970s (Friede, Busch and Bassen 2015). Since then, many studies have tried to study the topic using various methods (see: Dorfleitner, Kreuzer and Sparrer 2020; Kempf and Osthoff 2007; Renneboog, Ter Horst and Zhang 2008). Not only does the choice of methods differ, but there is also a wide range of results and conclusions that have been presented. These subsections aim to give the readers of this thesis an insight into previous studies. Further, the subsections strive to display the evolution of the literature.

2.4.1 The Relationship Between ESG-Investing and Performance

Kempf and Osthoff (2007) examine whether investors can get better performance from using different SRI screens. The conclusion is that buying companies with high socially responsible ratings and selling their counterparts positively affects performance. Nagy, Kassam and Lee (2016) also find a positive relationship between ESG and performance. This is tested by constructing two ESG portfolios. These results are analogous to the meta-study by Friede, Busch and Bassen (2015), which examines over 2000 empirical findings.
According to this meta-study, there is a nonnegative relationship between ESG and financial performance in approximately 90 percent of the examined literature. Although these studies find that superior financial performance is present, others such as (Brammer, Brooks and Pavelin 2006; Mănescu 2011) find evidence of underperformance for firms with a high level of sustainable awareness. In contrast, there is also a large part of the literature that neither can identify a positive or negative relationship. (see: Bauer, Koedijk and Otten 2005; Schröder 2007; Statman and Glushkov 2009).

In a confirmatory study, Revelli and Viviani (2014) conclude in their meta-analysis paper that SRI investing neither leads to under nor outperformance compared to conventional investing.

2.4.2 Explaining the Ambiguous Empirical Results

As shown above, the results regarding whether pursuing an ESG investing strategy leads to under- or outperformance relative to a non-ESG strategy are ambiguous. Pástor, Stambaugh and Taylor (2020) propose a theoretical model consistent with the heterogeneous empirical result. The model proposes that in equilibrium, "green" firms should underperform "brown" firms. This is because investors enjoy holding "green" firms. Investors prefer holding "green" firms because they derive a higher utility compared to holding "brown" firms.

A positive shock in investors' preferences for holding "green" firms, can lead to outperformance relative to their "brown" counterparts. Thus, the empirical results showing that ESG investing has outperformed during specific periods can be explained by an unexpected shift in preferences. Further, other studies that point towards ESG investing underperforming during specific periods are also consistent with the model since, in equilibrium, "green" firms will yield a lower return.

Another theoretical model that aims to explain the heterogeneous empirical results is presented by Pedersen, Fitzgibbons and Pomorski (2020). The model from Markowitz (1952) is extended by incorporating an ESG dimension. Three types of investors are considered in the model. The type-U investor is unaware of ESG scores and tries to maximize their unconditional mean-variance utility. In contrast, the type-A and type-M are not only trying to maximize their mean-variance utility. These investors also care about ESG scores. The type-A investor uses ESG scores to evaluate risk and expected returns, while the type-M investor prefers companies with high ESG scores. These companies are referred to as high ESG stocks in the paper. Depending on which investors that are dominating the economy, the expected returns on high ESG stocks will differ. If there are, for example, a large number of type-M investors, the price of high ESG stocks will be bid up. This will lead to lower expected returns on these stocks, all else equal. The prices will be bid up because type-M investors, which prefer high ESG stocks, are willing to receive lower returns due to holding high ESG stocks.

Dorfleitner, Kreuzer and Sparrer (2020) discussed how social pressure on institutional investors could lead to the exclusion of sin stocks. If institutional investors exclude these types of stocks, it may lead to higher returns. This is a result of a lower demand which lowers the price and, all else equal, raises the expected returns. In contrast, Nagy, Kassam and Lee (2016) present a rationale behind why ESG stocks instead could outperform relative to non-ESG stocks. By incorporating ESG into the operations, companies can avoid losses related to ESG. These losses are financial and therefore they affect the performance. Environmental fines or/and labor disputes are examples of what companies can avoid.

Barnett and Salomon (2006) also present a theoretical reasoning as to why ESG stocks can outperform non-ESG stocks. However, the explanation is different. The authors discuss how modern portfolio theory fails to account for the benefits that social screening offers. The reasoning is that even if the investment universe has constraints, the universe is superior. This is because the constrained universe includes stocks with a higher probability of beneficial performance in the future. Put differently, even if the investment universe is smaller, it contains “better” companies than the unconstrained universe. Capelle-Blancard and Monjon (2012) outline a different type of relationship. A U-shape relationship is presented by a combining modern portfolio theory and stakeholder theory. This relationship is between screening intensity and performance. The logic is as follows, as the screening becomes more intensive, picking better stocks will at least partially offset the cost of lower diversification.

2.4.3 Expected Returns and Realized Returns

Much of the literature in the area uses realized returns to measure performance (see: Auer 2014; Dorfleitner, Kreuzer and Sparrer 2020; Kempf and Osthoff 2007). Pástor, Stambaugh and Taylor (2022) emphasize distinguishing between returns ex-ante and ex-post. Building on their equilibrium model (see: Pástor, Stambaugh and Taylor 2020), the authors suggest that one should be careful about expecting higher returns on ESG investing going forward. One should be even more cautious if the expectations are based on previous outperformance. This is because the relationship between higher realized returns and expected returns is inverse. That higher realized returns are the result of low expected future is also documented by Fama and French (2002). Breaking down and distinguishing between expected and realized returns is not something novel. Elton (1999) highlights the importance of being aware of the distinction when interpreting results that are based on realized returns.

2.4.4 ESG-Investing and Investment Styles

The ability to categorize a portfolio based on its characteristics is helpful to determine how a particular portfolio fits the goal or criteria from an asset allocation perspective (Climent and Soriano 2011). Solutions for this are already well established in the fund industry. In 1992, Morningstar introduced its tool, the Morningstar Style Box. The tool was developed to help investment managers to identify the investment styles of certain funds (Kinnel 2002). This very same style box tool that is used to categorize and evaluate funds can also be used to evaluate portfolios consisting of individual stocks and bonds. For example, similarly, Climent and Soriano (2011) use a slightly modified approach in order to categorize their constructed portfolios along the style dimension. Ibikunle and Steffen (2015) also examine the different investment styles their constructed portfolios possess.

Their result suggests that ESG portfolios experience high levels of volatility, i.e. a high beta. One possible explanation is that ESG portfolios sometimes tend to exclude defensive traditional value stocks. According to Climent and Soriano (2011) these traditional value stocks can be found in sectors such as energy, chemical and basic industries. Furthermore, these types of companies often exhibit less market risk i.e. lower beta. According to the authors, this might explain the higher market sensitivity, which can be seen in some of these portfolios. Furthermore, another reason for the higher beta is the smaller investment universe available for these portfolios. This reasoning is in line with Climent and Soriano (2011) regarding that investing in a restricted universe of stocks implies a compromise of risk-reward optimization. As the investment universe is reduced, the potential for diversification is smaller than in an unconstrained portfolio.

The same authors also show that the examined portfolios with higher ESG scores are associated with small capitalization stocks. This is because the positive screening process has a tendency to exclude large capitalization stocks. However, others, such as Kaiser (2020), arrive at different results. It is instead shown that larger firms are associated with higher ESG scores. The positive relation between the two is attributed to increased shareholder pressure. Larger firms tend to have more pressure when it comes to sustainable responsibility. The literature is somewhat more in consensus regarding the tilt towards or away from growth stocks. According to the results of (Ibikunle and Steffen 2015; Kaiser 2020), funds composed of firms with high ESG scores are more exposed to growth stocks.

The given theoretical reasoning for this is that fast-growing companies tend to invest in innovative environmental and clean tech. However, some time-varying effects are seen by the former of the two. During the time period of 2004-2014, the funds with lower ESG-rated firms experience a growth stocks bias. The rationale given as to why this is the case is that these companies might need to adapt their investment strategies to maintain their returns. This is done by shifting towards emerging specialized high-growth enterprises in the fossil energy value chains. This may entail larger companies acquiring the shares or the direct operations of smaller firms operating in similar sectors of the economy.

The final factor examined by these papers is momentum. According to (Ibikunle and Steffen 2015; Kaiser 2020), portfolios with high scores are tilted away from momentum stocks. Thus, the two conclude that a positive relationship exists between portfolios with low ESG scores and momentum stocks. One potential explanation from Kaiser (2020) is that a relationship between momentum and media coverage seem to exist. Stocks that experience a negative trend in returns will try to increase their ESG performance to send positive signals to the market. Meanwhile, stocks that exhibit positive momentum are less concerned about their ESG performance.

2.4.5 The Non-Homogenous Views of ESG

As highlighted by Dorfleitner, Kreuzer and Sparrer (2020), the amount of ESG data available has increased during the last couple of decades. However, just as there are a lot of non-homogeneous views on the relationship between ESG and financial performance, there is divergence in how ESG is measured. The fact that ESG score providers use a mix of methods gives rise to discrepancies between the scores issued. This can result in the same company receiving different scores from different providers. Many challenges can therefore arise for actors in the financial markets. This is because ESG scores are one of the main tools for implementing ESG into investment strategies (Boffo and Patalano 2020). According to Ehlers et al. (2023), the most critical challenge is that investment managers may end up with different sets of portfolios while using the same investment strategy. This scenario can be a consequence of using different ESG score providers. The problem with different issuing parties using different assessment methods is also highlighted by Berg, Kölbel and Rigobon (2019). ESG ratings from six providers, Moody's, S&P Global, KLD, Refinitiv, MSCI and Sustainalytics are examined. By comparing ESG ratings to credit ratings, the inconsistency of the ESG ratings between the providers is illustrated.

In the study, it is shown that while the correlation between credit ratings usually lies around 0,99, the correlation for ESG scores ranges from 0,38 to 0,71. One potential explanation is that the definition of ESG is subjective. Furthermore, credit ratings are a concept that has matured over a more extended time period. This has resulted in a generally accepted framework compared to the relatively newer issuance of ESG scores.

As stated earlier, ESG investing lacks a formal definition which leads to variability in how it is implemented in practice. Auer (2014) discusses how the screening process has developed. Negative screening was the most common approach in the early years of responsible. However, this has changed and today, portfolios are constructed based on positive screens as well. It is proposed that a negative screening is superior to a positive screening approach. The theoretical rationale is that diversification worsens more with positive screening than negative screening. Lower diversification will raise the volatility of the portfolio, which in many cases is the metric used to measure risk. Therefore, the risk-adjusted returns of these portfolios will be lower, all else equal. Studies such as (Dorfleitner, Kreuzer and Sparrer 2020; Kempf and Osthoff 2007; Lioui and Tarelli 2022) construct portfolios following a positive screening. The results are portfolios based on the companies with the highest scores. Dorfleitner, Kreuzer and Sparrer (2020) use combined ESG scores and scores for each pillar. Furthermore, when looking at the individual pillars of ESG, Auer (2014) proposes that companies with outstanding governance should do better than less responsible firms.

2.4.6 Funds versus Portfolios of Individual Stocks

To examine the relationship between ESG investing and financial performance, previous studies are conducted using both equity funds and portfolios composed of individual stocks. For example, Ibikunle and Steffen (2015) compare the performance of “green” European mutual funds with their conventional and “black” peers.
Other examples are (Capelle-Blancard and Monjon 2012; Renneboog, Ter Horst and Zhang 2008), which compare socially responsible funds with their conventional peers. However, Kempf and Osthoff (2007) point out the drawbacks of using funds to examine the relationship between ESG and financial performance. The main critique is the failure to separate the manager's skill from the abovementioned relationship that these studies try to assess. A similar critique is expressed by (Statman and Glushkov 2009; Ziegler, Schröder and Rennings 2007), highlighting the difficulty of distinguishing between the performance attributable to management and the performance effects coming from sustainability.

Studies such as (Auer 2014; Dorfleitner, Kreuzer and Sparrer 2020) are instead following an approach of constructing portfolios composed of individual stocks.

2.4.7 How Performance is Defined and Measured

Since the relationship between ESG and financial performance is examined, a key question is how the literature measures performance. Different methods used to evaluate performance can lead to different results (Renneboog, Ter Horst and Zhang 2008). Studies such as Auer (2014) use the Sharpe ratio to measure performance. Other studies such as (Dorfleitner, Kreuzer and Sparrer 2020; Kempf and Osthoff 2007; Renneboog, Ter Horst and Zhang 2008; Statman and Glushkov 2009) use a multifactor model. Also, it is not uncommon for studies to combine different performance metrics (see: Statman and Glushkov 2009). Additionally, studies such as Ibikunle and Steffen (2015) use both CAPM and the Carhart four-factor model while referring to the criticism of CAPM.

Another dimension of the performance evaluation is highlighted by Dorfleitner, Kreuzer and Sparrer (2020). As the authors point out, it is not just a mixture of models and metrics that affects the heterogeneous empirical results but also a variety in how the literature defines financial performance. The performance can be defined from an accounting perspective focusing on measures such as earnings per share (EPS), return on equity (ROE) and net income. For example, Mardini (2022) runs regressions using both ROA and ROE to measure performance. Distinctively performance can be defined from a stock-market-oriented perspective. If the latter approach is used, metrics such as alpha, Tobin's Q and Sharpe ratio can be used. Kempf and Osthoff (2007) follow a stock-market-oriented technique using alpha as the performance measure.

Dorfleitner, Kreuzer and Sparrer (2020) suggest that a stock-market-oriented perspective is preferred if the analysis is conducted from an investor´s standpoint. This is because it better reflects investors´ ideas of the definition of future value. Further, Hillman and Keim (2001) express that a shortcoming of using accounting metrics is that they contain a more short-term perspective by nature. Additionally, it is explained that accounting metrics also can suffer from manipulations by management which is a limitation. Based on these arguments, Hillman and Keim (2001) also propose that market-based measures of performance are in favor of accounting-based measures.

2.4.8 The Time Period Examined

Further adding to the heterogeneity in the literature is the examined time period. The divergence can be seen both with regard to the length of the period being studied and if the period is split into smaller subperiods. Dorfleitner, Kreuzer and Sparrer (2020) divide the total time period into two equally long subperiods. A similar approach is used by Renneboog, Ter Horst and Zhang (2008). By doing so, it is possible to study the portfolios in different macroeconomic environments (Dorfleitner, Kreuzer and Sparrer 2020). Some literature focuses explicitly on the relationship between ESG stocks and market crashes. For example, Albuquerque et al. (2020) examine how stocks with high environmental and social scores behaved during the Covid-19 crash. The conclusion is that these stocks have higher returns and lower volatility compared to other stocks. Furthermore, Hoepner et al. (2016) provide evidence that companies can experience lower downside risk by addressing environmental, social and governance problems. Moreover, it is concluded that the environmental pillar has the highest impact concerning reducing downside risk.

Another dimension linked to the time period is brought up by Bennani et al. (2018). The paper describes how much of the literature has been done on long-term historical data. However, the tools for assesing corporate governance, environmental and social responsibility have changed. In addition, many of the tools are recent inventions.
This, in combination with the quick evolution of sustainable investing, leads to the conclusion that research on ESG asset pricing should only include more recent years. More specifically, it is reasoned that the time prior to the 2008 global financial crisis should be viewed with caution. By using recent data that benefit from higher confidence, one can avoid results not reflecting the current market structure.

3. Data

Sections 3 and 4 aim to describe the study's data collection process and the quantitative methods used in the thesis. Furthermore, the reader will also be presented with in-depth variable descriptions as well as an elaboration on the matter of robustness and statistical testing. A motivation behind all active choices made in the sections is presented. The intention is to increase the transparency surrounding the approaches used. By doing this, the reader an increased understanding of the later results.

3.1 Screening

3.1.1 ESG Scores: Combined and Individual

In line with Dorfleitner, Kreuzer and Sparrer (2020), a screening based on combined ESG scores and each pillar is done. This approach is applied since investors have heterogeneous perceptions and motives regarding ESG investing and how to implement it (see: section 2.4). Thus, investors might emphasize the importance of each pillar differently. The data for both the combined and individual scores are collected from one source, the Refinitiv Eikon database. Moreover, the database was chosen because of its transparent scoring methodology but also due to the availability of the platform. The scores from Refinitiv are based on publicly reported and available data such as annual reports, financial reports, CSR reports and press releases (Refinitiv 2022).

For the combined ESG score, the calculation is done by using the factors of ESG, with Environmental accounting for 37 percent, Social for 33 percent and Governance for 30 percent of the total score. Each factor is then built up of several subcategories (see: Table A2 in Appendix A). The weights of the subcategories are, in turn, determined by how many metrics each consists of. In total, the ESG score is made up of 630 metrics/data points, which brings about the overall ESG score. The combined score also incorporates the controversies score, which captures negative ESG scandals violating any of Refinitiv's 23 controversy topics. In short, the more scandals a company is involved in, the lower the score is (Refinitiv 2022). How one can measure ESG performance differs greatly (Berg, Kölbel and Rigobon 2019). By using a transparent provider such as Refinitiv, one can easily create their perception of Refinitivs approach and compare it to methods deployed by other issuers.

3.1.2 Geographical Focus and Screening Universe

It is not unusual that the US market is examined in previous literature (See: Climent and Soriano 2011; Dorfleitner, Kreuzer and Sparrer 2020). Furthermore, as of April 28 2023, US equity accounted for 60.52 percent of the MSCI ACWI All Cap Index (MSCI 2023). It is, therefore, possible to argue that the US is one of, if not the most, important equity markets in the world. As previously mentioned, the geographical focus for this thesis is, therefore, the US market. The literature review discusses how the usage of funds has some drawbacks when evaluating performance.

Based on this critique, the decision to use individual stocks was made. Following a similar approach as (Henriksson et al. 2019), the process of screening for firms is based on the stocks included in the S&P 500 Total Return Index (SPXTR) as of April 2023. A total return index differs from a price index in that it tracks both capital gains but also any cash distributions, e.g. dividends (S&P Dow Jones Indices 2023a).

3.1.2 Time period

One initial intention was to have as long a time period as possible. This was to be able to evaluate how the portfolios perform in different macroeconomic environments.
However, regard is also taken to the trade-off between examining a longer time period and the availability of data. ESG scores in Refinitiv's database are available from 2002 (Refinitiv 2022). Despite that, a limited number of companies have data on ESG scores in the first couple of years. In order to account for the disparity in the early years, the chosen time period over which the data is collected covers Jan 2007 - Dec 2022. This choice of time frame is also consistent with Bennani et al. (2018). As mentioned in the literature review, the authors suggest that an appropriate time period is from around 2008 and onwards.

3.1.3 Different Screening Approaches

In line with Auer (2014), both negative and positive screens are used. In addition, a third screening approach will be applied. The positive screening will be defined as when the companies with the highest scores are chosen. Instead, negative screening is not about choosing certain stocks but rather about excluding them. In this case, the negative screening will be defined as excluding certain industries. The third screening approach is similar to the positive screening, but instead of choosing the companies with the highest scores, the companies with the lowest scores are chosen. The positive screening approach uses the 80th percentile as the cut-off value. For the third screening approach, the 20th percentile will be used as the cut-off value. Normally Refinitiv updates its ESG scores on a yearly basis. According to Refinitiv (2022), the update is often done in conjunction with the company's annual reports. Therefore a separate screening will be done each year to account for the changing scores. If the screening was not made each year, it would imply an assumption the companies had the same scores throughout the whole period. We would argue that it is not a realistic assumption that a company has the same scores for 16 years and thus a yearly screening is done.

Negative screening by excluding certain industries is one of the more common ESG strategies (Boffo and Patalano 2020). We follow a similar industry exclusion approach as is done for the S&P 500 ESG Leaders index (see: Table A3 in Appendix A for the exhaustive list of the excluded industries). After all three screening approaches are employed, the final sample consists of nine sub-samples on which the portfolios will be based. The tenth portfolio will not be based on any screening approach. Instead, this portfolio will consist of all companies in S&P 500 as of April 2023. In section 5.1, a detailed review of all the portfolios will be presented.

3.1.4 Cleaning the Data

During the screening process, a few outliers were identified in terms of daily return anomalies. As mentioned by Blázquez-García et al. (2021) outliers can be of interest when using time-series data. Therefore, correction or adjustments for these anomalies were not justified since it could lead to a potential loss of information.

3.2 Estimation Model

To examine the research questions, the main estimation model will be a multifactor model. More specifically, in line with Dorfleitner, Kreuzer and Sparrer (2020), we will make use of the Carhart four-factor model. In addition, we follow a similar approach as Ibikunle and Steffen (2015) by using CAPM. By employing the Carhart Four-factor model, some of the shortcomings with CAPM, which have been discussed in section 2.3, are controlled for. To consider eventual time-varying aspects and trends, the sample is split into two equally long sub-periods. Thus, we follow the same approach as Dorfleitner, Kreuzer and Sparrer (2020). The first period is Dec 2007 - Dec 2014 and the second period is Jan 2015 - Dec 2022. All regressions will be conducted using the software program Stata.

The formal regression equation for CAPM is:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_{1,i}(R_{m,t} - R_{f,t}) + \varepsilon_{i,t} \tag{3}$$

Where:

$R_{i,t} = Return\ on\ portfolio\ i\ at\ time\ t$

$R_{f,t} = Risk\ free\ rate\ at\ time\ t$

$\alpha_i = CAPM\ alpha - the\ abnormal\ return\ on\ portfolio\ i$

$R_{m,t} - R_{f,t} = Excess\ return\ of\ the\ market\ at\ time\ t$

$\varepsilon_{i,t} = Error\ term\ for\ portfolio\ i\ at\ time\ t$

The formal regression equation for the Carhart Four-factor model is:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_{1,i}(R_{m,t} - R_{f,t}) + \beta_{2,i}R_{SMB,t} + \beta_{3,i}R_{HML,t} + \beta_{4,i}R_{MOM,t} + \varepsilon_{i,t} \quad (4)$$

Where:

$R_{i,t} = Return\ on\ portfolio\ i\ at\ time\ t$

$R_{f,t} = Risk\ free\ rate\ at\ time\ t$

$\alpha_i = Four - factor\ alpha - the\ abnormal\ return\ on\ portfolio\ i$

$R_{m,t} - R_{f,t} = Excess\ return\ of\ the\ market\ at\ time\ t$

$R_{SMB,t} = Fama - French's\ risk\ premium\ capturing\ size\ effect\ at\ time\ t$

$R_{HML,t} = Fama - French's\ risk\ premium\ capturing\ book - to - market\ effects\ at\ time\ t$

$R_{MOM,t} = Jegadeesh\ and\ Titman's\ risk\ premium\ capturing\ momentum\ effects\ at\ time\ t$

$\varepsilon_{i,t} = Error\ term\ for\ portfolio\ i\ at\ time\ t$

3.3 Variables

The portfolio's excess returns will be used as the dependent variable in the regressions. For the Carhart model, excess market return ($R_m - R_f$), small minus big (SMB), high minus low (HML) and momentum (MOM) will be used as independent variables. Data for the four independent variables are collected from the Kenneth R. French database. The rationale behind this decision was the favorable accessibility to daily data.

All the data from the database includes both dividends and capital gains. Further, the data are not continuously compounded and are in US dollars. The portfolio used as a proxy for the market portfolio is a value-weighted portfolio. It consists of all shares listed on NYSE, AMEX and NASDAQ. Additionally, all shares have a CRSP code of 10 or 11 which means that only common shares are used. The Kenneth R. French database uses the one-month treasury bill from Ibbotson Associates as the risk-free rate. To calculate the excess market return, the risk-free rate will be subtracted from the market portfolio returns (French 2023).

The formal definition of a value-weighted portfolio:

$$\frac{EqV_{i,t}}{\Sigma_{j=1}^{Nt} EqV_{i,t}} \quad (5)$$

Where:

$EqV_{i,t} = Equity\ value\ of\ firm\ i\ at\ time\ t$

The median size of NYSE stocks is used to group all stocks on NYSE, AMEX and NASDAQ into one of the categories: big or small. All stocks are then sorted one more time based on their book-to-market ratio (B/M). The stocks are divided into one of three categories: low, medium or high. The bottom 30 percent is considered low, the middle 40 percent is medium and the top 30 percent is high. Stocks included in the category low can also be called growth stocks, while the stocks included in the category high can be referred to as value stocks. After the sorting six different value-weighted portfolios are created to calculate SMB, HML and MOM. The portfolios are referred to as: Small/Low, Small/Medium, Small/High, Big/Low, Big/Medium and Big/High (Bodie, Kane and Marcus 2020; French 2023)

The SMB factor is constructed by taking the average return on the three small stock portfolios and subtracting the average return on the three big stock portfolios.
Formally the equation for the SMB factor is:

$$SMB = \frac{(Small\ Value + Small\ Medium + Small\ Growth)}{3} - \frac{(Big\ Value + Big\ Medium + Big\ Growth)}{3} \quad (6)$$

Taking the average return on the two value portfolios and subtracting the average return on the two growth portfolios will yield the HML factor. Formally the equation for the HML factor is:

$$HML = \frac{(Small\ Value + Big\ Value)}{2} - \frac{(Small\ Growth + Big\ Growth)}{2} \quad (7)$$

Jegadeesh and Titman (1993) discovered the phenomenon that both good and bad performance seem to persist over a time period of several months.

This phenomenon is referred to as a sort of momentum. Carhart (1997) added the momentum factor to the Fama-French Three-factor model. The modified model with momentum is often labeled the Carhart Four-factor model. Portfolios constructed on momentum are also value-weighted and built based on size and prior returns in the last 2 to 12 months. Based on the monthly returns in the past 2 to 12 months, the stocks will be grouped into the category high or low. The low category consists of the lowest 30 percent and the high category consists of the highest 30 percent. The top and bottom 30 percent of the stocks traded on NYSE will be used as the criterion for sorting the stocks on NYSE, AMEX and NASDAQ into either the high or low category. Calculations for the momentum factor are computed as the average return for the two high portfolios minus the average return for the two low portfolios (Bodie, Kane and Marcus 2020; French 2023).

Formally the equation for the MOM factor is:

$$MOM = \frac{(Small\ High + Big\ High)}{2} - \frac{(Small\ Low + Big\ Low)}{2} \tag{8}$$

3.4 Potential Biases

Look-ahead bias can be described as when information that was not available at the time of the simulation is still used in the simulation. It is common that this skews the results to look better than they should (Daniel, Sornette and Woehrmann 2009). Refinitiv Eikon has been criticized for changing the ESG scores retroactively and thus causing look-ahead bias (Berg, Kölbel and Rigobon 2019). According to Refinitiv (2022), the scores issued longer than five years ago are not changed as they are marked as definitive. However, Berg, Kölbel and Rigobon (2019) show that Refinitiv can change scores even if they are marked as definitive. In March 2020 Refinitiv changed scores back to 2011, which illustrates the eventual problems with look-ahead bias. Since even the definitive scores cannot be considered definitive, we want to display the possibility that look-ahead bias is present in the thesis. Another potential bias that can exist in the thesis is the survivorship bias. Brown et al. (1992) present how survivorship bias can be contained in performance studies. Using data that only includes the survivors can cause the performance numbers to be biased. Since the base for the portfolios are the companies included in the S&P 500 as of April 2023 it is possible that survivorship bias is present in the thesis.

4. Methodology

This section aims to provide the reader with information regarding portfolio construction. Furthermore, as ratios will be used as a complement to the estimation models in order to evaluate the portfolios, this section also aims to describe these ratios. Important assumptions regarding transaction costs will also be highlighted. Lastly, we explain the statistical tests that are performed on the data.

4.1 Portfolio Construction

Based on the subsamples from the screening, ten different portfolios are constructed. Following the approach of Kaiser (2020), all the portfolios will be value-weighted. All the portfolio construction and data gathering have been made in Refinitiv Eikon. In the cases where data for a company is not available for the whole year, the company is excluded for that particular year. One dataset on the daily excess returns is constructed for each of the ten portfolios. Studies such as (Dorfleitner, Kreuzer and Sparrer 2020; Drei et al. 2019; Kempf and Osthoff 2007) construct long-short portfolios. Alternatively, Kaiser (2020) constructs long-only portfolios. The rationale behind using long-only portfolios is that a lot of investors, both retail and institutional, are not permitted to short-selling stocks. For example, UCITS funds are prohibited from taking direct short positions (Jennings 2008). Therefore the result of using long-only portfolios will be more relatable to investors, which is pointed out by Kaiser (2020). Based on this reasoning, this thesis will use long-only portfolios.

Each of the ten constructed portfolios is defined below:

Portfolio 1: Sector Exclusion
A value-weighted portfolio excluding industry-specific companies in line with the method used in the S&P 500 ESG Leader index.

Portfolio 2: Best in Class ESG
A value-weighted portfolio consisting of the top 20 percent of the companies with the highest ESG scores in the S&P 500 for each specific year.

Portfolio 3: Best in Class Environmental

A value-weighted portfolio consisting of the top 20 percent of the companies with the highest environmental scores in the S&P 500 for each specific year.

Portfolio 4: Best in Class Social

A value-weighted portfolio consisting of the top 20 percent of the companies with the highest social scores in the S&P 500 for each specific year.

Portfolio 5: Best in Class Governance

A value-weighted portfolio consisting of the top 20 percent of the companies with the highest governance scores in the S&P 500 for each specific year.

Portfolio 6: Conventional

A value-weighted portfolio consisting of all active companies included in the S&P 500 as of April 2023.

Portfolio 7: Worst in Class ESG

A value-weighted portfolio consisting of the bottom 20 percent of the companies with the lowest ESG score in the S&P 500 for each specific year.

Portfolio 8: Worst in Class Environmental

A value-weighted portfolio consisting of the bottom 20 percent of the companies with the lowest environmental score in the S&P 500 for each specific year.

Portfolio 9: Worst in Class Social

A value-weighted portfolio consisting of the bottom 20 percent of the companies with the lowest social score in the S&P 500 for each specific year.

Portfolio 10: Worst in Class Governance

A value-weighted portfolio consisting of the bottom 20 percent of the companies with the lowest governance score in the S&P 500 for each specific year.

4.2 Ratios for Performance Evaluation

Complementing the estimation models, ratios will be used to evaluate the portfolios. More specifically, the Sharpe ratio and Treynor ratio will be calculated. As previously mentioned, these measures belong to the so-called stock market-oriented perspective (see: Dorfleitner, Kreuzer and Sparrer 2020; Hillman and Keim, 2001). This thesis, therefore, follows a similar approach as Auer (2014). The Sharpe ratio is a measure of risk-adjusted performance since it measures the average excess return relative to its risk. Risk is measured as the standard deviation in the context of the Sharpe ratio. Therefore, the Sharpe ratio presents a procedure of returns per unit of risk (Opdyke 2007).

The formal equation for the Sharpe ratio is:

$$\frac{R_i - R_f}{\sigma_i} \qquad (9)$$

Where:

$R_i = Average\ return\ on\ portfolio\ i$

$R_f = Average\ risk - free\ rate$

$\sigma_i = Standard\ deviation\ of\ the\ excess\ returns\ on\ portfolio\ i$

Opdyke (2007) mentions that remarkably little research has focused on the statistical properties of the Sharpe ratio. This is in comparison to the research focusing on the utility of the Sharpe ratio for evaluating performance. Nevertheless, there is some research focusing on the statistical properties of the Sharpe ratio. Christie (2005) presents some of the limitations concerning using the Sharpe ratio as a measure of performance. One of the limitations stated is that Sharpe ratios are not appropriate when returns are non-normal. Ledoit and Wolf (2008) also bring up some shortcomings with the Sharpe ratio.
The discussion is about how hypothesis testing with the Sharpe ratio is not valid when returns are from a time series and non-normal.

As a result of the discussion regarding the statistical properties of the Sharpe ratio the approach suggested by Wright et al. (2014) is used. This approach does not require the returns to be identically and independently distributed (iid). The technique is applied in previous research, such as Amon et al. (2021). This paper is in the area of ESG investing, which further supports the idea of using the proposed formula in this thesis.

The formal equation for the Sharpe ratio suggested by Wright et al. (2014) is:

$$\widehat{SR}_i = \frac{[1/n]\Sigma_{t=1}^{n} R_{i,t} - R_{f,i}}{\sqrt{[1/n]\Sigma_{t=1}^{n}(R_{i,t} - R_{f,i})^2 - ([1/n]\Sigma_{i=1}^{n} R_{i,t} - R_{f,i})^2}} \tag{10}$$

The Sharpe ratio defines risk as the standard deviation. How to define risk is a discussion with different views. This has given rise to ratios defining risk in other ways. The Treynor ratio is one of these ratios that has another definition of risk. Instead of using the standard deviation as the definition of risk, the Treynor ratio uses beta. The rationale behind this is linked to the theory of CAPM. If it is only the systematic risk that is priced, it can be argued that the Treynor ratio is a better measure for performance evaluation. (Bodie, Kane and Marcus 2020). To make the results more robust, the Treynor ratio will thus be used in addition to the Sharpe ratio.

The formal equation for the Treynor Ratio is:

$$\frac{R_i - R_f}{\beta_i} \tag{11}$$

Where:

$R_i = \ Average\ return\ on\ portfolio\ i$

$R_f = Average\ risk\ -\ free\ rate$

$\beta_i = Standard\ deviation\ of\ the\ excess\ returns\ on\ portfolio\ i$

4.3 Transaction Costs

The assumption is made that all portfolios have similar transaction costs, and therefore, no adjustments are made when evaluating the portfolios.
The reason is that long-only portfolios are used. If long-short portfolios were constructed, there would be a motive for incorporating transaction costs. This is since shorting stocks can be costly (Jones and Lamont 2002). Moreover, eight out of ten portfolios have approximately the same turnover. All the portfolios except the Sector Exclusion portfolio and the Conventional portfolio are rebalanced every year. Even though there exist inconsistencies in the rebalancing frequency, we considered the differences to be negligible in this context. This is in line with Auer (2014), that demonstrates that transaction costs tend to be negligible if portfolios are rebalanced with a low frequency.

4.4 Statistical Test and Robustness

Since time-series data is used, several tests are conducted to ensure that the assumption of stationarity and autocorrelation holds. In addition, tests for the necessary OLS assumptions will be employed (Stock and Watson 2019). The Augmented Dickey-Fuller test is employed. This is a unit root test checking the stationarity assumption (Mushtaq 2011). A Breusch-Godfrey test will be applied to test if autocorrelation is present in the time series (Uyanto 2020). All tests are conducted in the software program Stata and presented in Table B1-B6 in Appendix B.

5. Results and Analysis

This section aims to answer the thesis's two research questions by presenting and analyzing the main findings. Firstly, descriptive statistics will be provided followed by an examination of portfolio performance and style. The results are examined within the context of the previous literature and theories. Lastly, a robustness check will be presented.

5.1 Descriptive Statistics

This section is begun by providing a statistical summary of the data used. The following three tables all display descriptive statistics, including mean, standard deviation, minimum and maximum value. Table I presents these statistics for the independent variables, while Tables II and III demonstrate it for the portfolios. In addition, skewness and kurtosis are provided for all the portfolios. The values are calculated using the daily data from the whole sample period of Jan 2007 to Dec 2022. Furthermore, Tables C1 and C2 in Appendix C show the values for the two subperiods 2007-2014 and 2015-2022.

Table I – Descriptive Statistics Independent Variables

Variables	Mkt-Rf	SMB	HML	MOM
Mean	0.03817	0.00125	-0.00454	0.00569
Median	0.08	0.00	-0.03	0.07
St. Dev	1.31259	0.62816	0.85497	1.12277
Min	-12.00	-3.79	-5.00	-14.37
Max	11.35	5.48	6.74	7.12

Notes to Table I: The mean and standard deviation have been rounded to five decimals, while the median, min and max have been rounded to two decimals. The number of observations for all independent variables is 4028.

Table II – Descriptive Statistics ESG Portfolios

Variables	Sector Exclusion	Best in Class ESG	Best in Class Environmental	Best in Class Social	Best in Class Governance
Mean	0.05519	0.04412	0.04246	0.04096	0.05138
Median	0.11081	0.09029	0.07896	0.08131	0.09832
St. Dev	1.30917	1.24193	1.24992	1.25546	1.37978
Min	-12.52766	-11.67653	-11.31580	-10.43327	-12.18762
Max	10.28456	10.22028	10.83815	11.09206	11.65520
Skewness	-0.32563	-0.25762	-0.30489	-0.23655	-0.27562
Kurtosis	12.49735	13.96334	14.01525	13.46263	13.18313

Notes to Table II: The table displays descriptive statistics for the excess returns of each ESG portfolio. All numbers have been rounded to five decimals. The number of observations for all the portfolio's excess returns is 4028.

Table III – Descriptive Statistics Non-ESG Portfolios

Variables	Conventional	Worst in Class ESG	Worst in Class Environmental	Worst in Class Social	Worst in Class Governance
Mean	0.05368	0.06665	0.06466	0.06183	0.05614
Median	0.10244	0.10455	0.10527	0.10056	0.11272
St. Dev	1.29645	1.37318	1.34296	1.36006	1.43447
Min	-12.54160	-13.67849	-14.67725	-14.36198	-12.13761
Max	10.31037	13.34239	13.10246	13.36991	13.76405
Skewness	-0.35255	-0.16673	-0.21829	-0.17604	-0.05689
Kurtosis	12.99198	13.80310	15.36413	15.07917	13.33126

Notes to Table III: The table displays descriptive statistics for the excess returns of each non-ESG portfolio. All numbers have been rounded to five decimals. The number of observations for all the portfolio's excess returns is 4028.

Table II illustrates that the Sector Exclusion portfolio has the highest average excess return out of the ESG portfolios.

However, the Sector Exclusion portfolio also demonstrates a high level of dispersion relative to the best in class portfolios. Only the Best in Class Governance portfolio shows more fluctuations in the data. Examining the results in Table III, one can see that the Worst in Class ESG portfolio yields the highest average excess return. Focusing on the dispersion, the Worst in Class Governance portfolio has the highest standard deviation.

5.2 Research Question I - How do the risk-adjusted returns differ between the constructed portfolios?

To answer the first research question the structure is the following. To start with, the results from the Carhart model are presented and analyzed. Continuing, the findings obtained applying the Sharpe- and Treynor ratios are demonstrated and assessed.

As mentioned in the literature review, alpha is a common measure to evaluate risk-adjusted returns. The theory and method sections have also examined the rationale behind not only using a one-factor model. Based on that reasoning, the section will present the results obtained from the Carhart model. The results from using CAPM are presented in Table C3-C8 in Appendix C. When interpreting and analyzing all the results, emphasis will be on the portfolios displaying a statistically significant alpha on a one- or five percent level.

Table IV: ESG Portfolios Carhart Four-Factor Model

Variables	**Sector Exclusion**	**Best in Class ESG**	**Best in Class Environmental**	**Best in Class Social**	**Best in Class Governance**
Alpha	0.01808***	0.00866**	0.00698*	0.00450	0.01276**
	(0.00425)	(0.00441)	(0.00385)	(0.00376)	(0.00566)
Mkt-RF	0.96998***	0.93065***	0.93841***	0.95394***	1.01329***
	(0.00856)	(0.00577)	(0.00532)	(0.00528)	(0.00734)
SMB	0.06506***	-0.08030***	-0.13639***	-0.15013***	-0.01059
	(0.01508)	(0.01199)	(0.00879)	(0.00787)	(0.01406)
HML	-0.00550	0.02301***	0.06386***	-0.05437***	0.02115**
	(0.00883)	(0.00958)	(0.00821)	(0.00812)	(0.01061)
MOM	-0.00402	0.02595***	0.02023***	-0.00302	0.00125
	(0.00668)	(0.00652)	(0.00739)	(0.00614)	(0.00936)
R^2	0.9573	0.9492	0.9622	0.9637	0.9325

Notes to Table IV: Robust standard errors are displayed in parentheses. All coefficients and robust standard errors are rounded to five decimals. All R^2 are rounded to four decimals. The number of observations is 4028 for all the portfolios.

* Significant at a 10 percent level, ** Significant at a 5 percent level, *** Significant at a 1 percent level.

Table V: Non-ESG portfolios Carhart Four-Factor Model

Variables	Conventional	Worst in Class ESG	Worst in Class Environmental	Worst in Class Social	Worst in Class Governance
Alpha	0.01680***	0.02990***	0.02988***	0.02644***	0.01622**
	(0.00410)	(0.00873)	(0.00887)	(0.00883)	(0.00680)
Mkt-RF	0.96430***	0.95181***	0.91357***	0.92371***	1.03948***
	(0.00736)	(0.01921)	(0.01939)	(0.01941)	(0.00923)
SMB	0.04799***	0.11731***	0.17209	0.17004***	0.04894***
	(0.01281)	(0.03390)	(0.03242)	(0.03220)	(0.01634)
HML	0.00609	-0.06365***	-0.01769	0.00740	-0.07563***
	(0.00757)	(0.01735)	(0.01783)	(0.01775)	(0.01233)
MOM	0.00645	-0.00227	-0.00147	-0.00753	-0.02840***
	(0.00576)	(0.01328)	(0.01455)	(0.01393)	(0.01140)
R^2	0.9595	0.8352	0.8241	0.8289	0.9094

Notes to Table V: Robust standard errors are displayed in parentheses. All coefficients and robust standard errors are rounded to five decimals. All R^2 are rounded to four decimals. The number of observations is 4028 for all the portfolios.
* Significant at a 10 percent level, ** Significant at a 5 percent level, *** Significant at a 1 percent level.

Tables IV and V display that all alphas are statistically significant except for one portfolio. This is the Best in Class Social portfolio. Furthermore, the alpha for the Best in Class Environmental portfolio is significant at a ten percent level. It can be concluded from the results that overall the non-ESG portfolios have higher alphas than the ESG portfolios. These results are consistent with (Brammer, Brooks and Pavelin 2006; Mănescu 2011), which display underperformance for firms with a high level of sustainable awareness. Furthermore, that investing in “brown” stocks will yield a higher return than investing in “green” stocks is in line with the theory proposed by Pástor, Stambaugh and Taylor (2020). As described in the literature review, in equilibrium, investors need to be compensated for holding “brown” stocks. If preferences for “green” stocks unexpectedly shift, this can result in higher realized returns for “green” stocks. The results imply that according to this theory, the preferences have not shifted unexpectedly during 2007-2022, or at least not enough, to make “green” stocks outperform “brown” stocks.

Moreover, the results are also consistent with Dorfleitner, Kreuzer and Sparrer (2020), who discuss the effect of social pressure on institutional investors. A potential explanation behind the findings can therefore be that institutional investors have experienced increased pressure. Many institutional investors explicitly have expressed preferences for companies with high ESG ratings (Lioui and Tarelli 2022).

Whether this is due to pressure or not will not be speculated on. However, the implications of institutional investors excluding companies that are not considered to meet the ESG criteria are still the same. It will lead to lower demand which in turn lowers the price and thus raises the returns, all else equal. Potentially, this is what is observed in the results.

The phenomenon that non-ESG portfolios have a higher alpha than ESG portfolios is not in line with the findings from Kempf and Osthoff (2007). As described in the literature review, one of the main conclusions from that paper is that investors can earn positive abnormal returns by following an ESG strategy. By investing in stocks with high SRI ratings while short-selling stocks with low SRI ratings, investors can earn positive abnormal returns. Nagy, Kassam and Lee (2016) also show that by overweighting companies with higher ESG scores, investors are able to achieve higher returns. As mentioned in the literature review, an explanation provided is linked to the avoidance of, for example, environmental fines or/and labor disputes. The results from the regressions run in this thesis do not support that rationale. Even if it is the case that companies can perform better as a consequence of avoiding losses related to ESG, it does not show in the results. Or at least, it is not enough to compensate for the eventual costs of pursuing an ESG strategy.

Both the results from (Kempf and Osthoff 2007; Nagy, Kassam and Lee 2016) are in opposition to the outcome obtained in this thesis. Once again, this shows one of the key concerns in the research on ESG investing. This concern has been emphasized a lot throughout the thesis and it is the heterogeneity. The results, that a non-ESG strategy overall delivers a higher alpha relative to an ESG strategy, are thus both consistent and inconsistent with previous literature.

In accordance with the model from Pedersen, Fitzgibbons and Pomorski (2020), one interpretation of the regressions is that it seems like type-M investors have dominated the market. Again this is compatible with the view that institutional investors have a preference for firms with high ESG ratings (Loui and Tarelli 2022). Since these investors have a preference for companies with higher ESG scores, they are willing to pay a higher price and thus, all else equal, get a lower return. If type-M investors are dominating the market, ESG stocks will underperform relative to non-ESG stocks, which are displayed in the results.

As discussed previously, the literature uses different types of screens when constructing ESG portfolios. According to the results, pursuing an ESG approach using negative screening is superior to positive screening. Table IV demonstrates that the Sector Exclusion portfolio has the highest alpha out of the ESG portfolios. These results are in line with Auer (2014), which proposes that SRI investors should apply negative screens rather than positive screens. Still, the outcome is a bit puzzling. According to the reasoning, the Conventional portfolio should perform better than the Sector Exclusion portfolio. This is since the Sector Exclusion portfolio has a lower diversification than the Conventional portfolio.

A potential reason behind these findings is that the “all else equal” does not hold. There seem to be other factors that make the Sector Exclusion portfolio have a higher alpha than the Conventional portfolio. The rationale brought up by Barnett and Salomon (2006) may help to explain the results. Potentially the Sector Exclusion portfolio excludes the worst companies without lowering diversification too much. The Sector Exclusion portfolio thus gets some of the benefits and some of the costs. However, it seems that the net effect is positive. This is also, to some extent, analogous to the U-shaped relationship presented by Capelle-Blancard and Monjon (2012).

The U-shape relationship is consistent with the fact that the Sector Exclusion portfolio performs relatively well. Despite that, it is not fully compatible with the overall results. If the U-shape relationship would hold, it would imply that the best in class portfolios should perform more similarly to the Sector Exclusion portfolio. Our results, therefore, do not indicate that the lower diversification is offset by only investing in the “best companies”. Instead, the results suggest that as more stocks are excluded, it will have a negative effect, at least after a certain point.

The results also reveal that the Best in Class Governance portfolio has a higher alpha compared to the other best in class portfolios. These findings are partially aligned with the reasoning about the governance pillar from Auer (2014). Additionally, the Worst in Class Governance portfolio has an inferior alpha compared to the other worst in class portfolios. The results thus point towards that investing in companies with a bad governance score has a negative effect on performance. It is important to note that the results are just partly in line with Auer (2014). A puzzling dimension is that the Worst in Class Governance portfolio still has a higher alpha than the Best in Class Governance portfolio.

Hence, it appears like there are additional aspects that affect the results. One potential explanation can be that companies with low governance scores also have low environmental, social and ESG scores or at least one of the others. As has been discussed previously, in order for investors to hold these stocks they need to be compensated. Therefore, it can be the case that there are forces that go in opposite directions. Firstly, firms with bad governance tend to perform worse than firms with good governance. Secondly, in order to even hold these stocks, investors require a higher compensation. In the results, the higher demanded compensation seems to be the dominating effect.

As outlined in section 4.2, the aim is to account for eventual time-varying aspects. By doing this, it is possible to see how the portfolios behave in different macroeconomic environments. The regressions on the subperiods are presented in Tables VI-IX.

Table VI: ESG Portfolios 2007-2014 Carhart Four-Factor Model

Variables	**Sector Exclusion**	**Best in Class ESG**	**Best in Class Environmental**	**Best in Class Social**	**Best in Class Governance**
	2007-2014	2007-2014	2007-2014	2007-2014	2007-2014
Alpha	0.02144***	0.00761	0.01028**	0.00337	0.00504
	(0.00655)	(0.00524)	(0.00482)	(0.00472)	(0.00783)
Mkr-RF	0.93343***	0.94886***	0.96927***	0.96954***	1.02919***
	(0.01209)	(0.00803)	(0.00604)	(0.00614)	(0.01016)
SMB	0.15754***	-0.08482***	-0.12744***	-0.14213***	0.02207*
	(0.02264)	(0.01328)	(0.01181)	(0.01021)	(0.02137)
HML	0.06086***	-0.08971***	-0.00345	-0.03382***	-0.02809**
	(0.01999)	(0.01405)	(0.01384)	(0.01328)	(0.02046)
MOM	-0.04034***	0.00209	-0.00504	-0.01573***	-0.12042***
	(0.01087)	(0.00690)	(0.00787)	(0.00726)	(0.01284)
R^2	0.9559	0.9680	0.9753	0.9757	0.9481

Notes to Table VI: Robust standard errors are displayed in parentheses. All coefficients and robust standard errors are rounded to five decimals. All R^2 are rounded to four decimals. The number of observations is 2014 for all the portfolios.

* Significant at a 10 percent level, ** Significant at a 5 percent level, *** Significant at a 1 percent level.

Table VII: Non-ESG Portfolios 2007-2014 Carhart Four-Factor Model

Variables	Conventional	Worst in Class ESG	Worst in Class Environmental	Worst in Class Social	Worst in Class Governance
	2007-2014	2007-2014	2007-2014	2007-2014	2007-2014
Alpha	0.01958***	0.03479**	0.03508**	0.03150**	0.02252**
	(0.00659)	(0.01463)	(0.01407)	(0.01416)	(0.01106)
Mkr-RF	0.94452***	0.86577***	0.8261104***	0.84961***	1.05757***
	(0.01052)	(0.02993)	(0.02715)	(0.02859)	(0.01279)
SMB	0.11594***	0.18825***	0.23655***	0.22398***	0.08720***
	(0.01986)	(0.05516)	(0.04950)	(0.05206)	(0.02585)
HML	0.03027***	0.05408**	0.07525***	0.06340**	-0.01845
	(0.01768)	(0.04783)	(0.04365)	(0.04610)	(0.02785)
MOM	-0.01785**	-0.01582	-0.04554***	-0.02881*	-0.06760***
	(0.00979)	(0.02632)	(0.02449)	(0.02554)	(0.02057)
R^2	0.9546	0.7862	0.7935	0.7951	0.9038

Notes to Table VII: Robust standard errors are displayed in parentheses. All coefficients and robust standard errors are rounded to five decimals. All R^2 are rounded to four decimals. The number of observations is 2014 for all the portfolios.
* Significant at a 10 percent level, ** Significant at a 5 percent level, *** Significant at a 1 percent level.

During the first subperiod, it still seems like the non-ESG portfolios overall perform better than the ESG portfolios. The period between 2007-2014 includes both the financial crisis and the euro crisis. As previously mentioned, studies such as Albuquerque et al. (2020) suggest that sustainable stocks can perform better during crises. Similarly, Hoepner et al. (2016) present that ESG engagement can reduce companies' downside risk. Despite that, the results from the regressions point towards the opposite. All the worst in class portfolios perform strongly during the first period. Moreover, all alphas are also significant. Nonetheless, one should be careful when interpreting the results and drawing conclusions.

The period between 2007-2014 includes not only the financial- and euro crisis but also at least some of the recovery period afterward. It can therefore be the case that the ESG portfolios perform better during the actual crises. However, when extending the period, they perform worse. This is not a topic that is specifically examined in the thesis. Nevertheless, the results suggest that there is no additional gain during crises by holding ESG portfolios. The only part of the results that are in some sense consistent with the literature is that the Best in Class Environmental portfolio performs better than the other best in class portfolios.

This is in line with the findings from Hoepner et al. (2016), suggesting that the environmental pillar has the highest impact on downside risk. Still, all the best in class portfolios seem to be inferior compared to the worst in class portfolios which are in opposition to the literature.

Table VIII: ESG Portfolios 2015-2022 Carhart Four-Factor Model

Variables	**Sector Exclusion**	**Best in Class ESG**	**Best in Class Environmental**	**Best in Class Social**	**Best in Class Governance**
	2015-2022	2015-2022	2015-2022	2015-2022	2015-2022
Alpha	0.01505***	0.00823	0.00322	0.00638	0.02072***
	(0.00491)	(0.00685)	(0.00583)	(0.00566)	(0.00715)
Mkr-RF	0.98089***	0.93775***	0.91115***	0.92031***	0.96296***
	(0.00936)	(0.00920)	(0.00825)	(0.00714)	(0.00906)
SMB	0.00272	-0.09123***	-0.14539***	-0.14316***	0.00811
	(0.01459)	(0.01871)	(0.01158)	(0.01099)	(0.01544)
HML	-0.03683***	0.07379***	0.07835***	-0.08036***	0.00611
	(0.00880)	(0.01221)	(0.00971)	(0.00933)	(0.01018)
MOM	0.02064***	-0.02564	0.02939***	0.01179**	0.08665***
	(0.00662)	(0.00998)	(0.01100)	(0.00759)	(0.00816)
R^2	0.9666	0.9303	0.9458	0.9500	0.9285

Notes to Table VIII: Robust standard errors are displayed in parentheses. All coefficients and robust standard errors are rounded to five decimals. All R^2 are rounded to four decimals. The number of observations is 2014 for all the portfolios.

* Significant at a 10 percent level, ** Significant at a 5 percent level, *** Significant at a 1 percent level.

Table IX: Non-ESG Portfolios 2015-2022 Carhart Four-Factor

Variables	Conventional	Worst in Class ESG	Worst in Class Environmental	Worst in Class Social	Worst in Class Governance
	2015-2022	2015-2022	2015-2022	2015-2022	2015-2022
Alpha	0.01391***	0.02527***	0.02474**	0.02091**	0.01145
	(0.00468)	(0.00900)	(0.01031)	(0.01014)	(0.00725)
Mkr-RF	0.97253***	1.02038***	0.98150***	0.99569***	0.97884***
	(0.00962)	(0.01784)	(0.02218)	(0.02231)	(0.00913)
SMB	-0.00449	0.06994***	0.13964***	0.12881***	0.04952***
	(0.01333)	(0.03291)	(0.03528)	(0.03249)	(0.01464)
HML	-0.00569	-0.08843***	-0.03470***	0.01127	-0.13389***
	(0.00843)	(0.01460)	(0.01787)	(0.01735)	(0.01026)
MOM	0.01855***	0.01197	0.03343***	0.00462	0.01084
	(0.00624)	(0.01191)	(0.01494)	(0.01365)	(0.00822)
R^2	0.9689	0.9041	0.8722	0.8769	0.9325

Notes to Table IX: Robust standard errors are displayed in parentheses. All coefficients and robust standard errors are rounded to five decimals. All R^2 are rounded to four decimals. The number of observations is 2014 for all the portfolios.
* Significant at a 10 percent level, ** Significant at a 5 percent level, *** Significant at a 1 percent level.

The results show that the non-ESG portfolios overall display higher alphas than the ESG portfolios. This has been the case, independent of which time period the regressions are conducted. The results, therefore, seem to be unvarying with regard to specific time periods. In contrast, a portfolio that behaves very differently when the sample is split is the Best in Class Governance portfolio. During the first period, it has one of the lowest alphas. Furthermore, the alpha is not statistically significant. On the contrary, the portfolio has the highest alpha out of the ESG portfolios during the second period. Thus, the results reveal that there is a time-varying aspect related to the Best in Class Governance portfolio. The results are, therefore, both consistent and inconsistent with Auer (2014). This time-varying result can be considered enigmatic. None of the theoretical models or literature examined have a logical rationale for these results.

We now proceed to measure the risk-adjusted returns with the financial ratios, starting off by presenting the Sharpe ratios. Tables XIII and XIV illustrate the Sharpe ratios during the whole sample period for each of the constructed portfolios.

Table X: Sharpe Ratios ESG Portfolios

Variables	Sector Exclusion	Best in Class ESG	Best in Class Environmental	Best in Class Social	Best in Class Governance
Sharpe Ratio	0.04216	0.03553	0.03397	0.03263	0.03725
t-statistic	0.44767	0.22649	0.17109	0.12387	0.27185
p-value	0.65440	0.82983	0.86416	0.90142	0.78575

Notes to Table X: The one-month treasury bill from Ibbotson Associates has been used as the risk-free rate. When computing the t-statistic all portfolios have been compared to the market portfolio. The portfolio used as a proxy for the market portfolio is the same as for the regressions. This a value-weighted portfolio consisting of all shares listed on NYSE, AMEX or NASDAQ. All numbers have been rounded to five decimals. * Significant at the 10 percent level, ** Significant at the 5 percent level, *** Significant at the 1 percent level.

The formula used for the t-test is: $t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$

Table XI: Sharpe Ratios non-ESG Portfolios

Variables	Conventional	Worst in Class ESG	Worst in Class Environmental	Worst in Class Social	Worst in Class Governance
Sharpe Ratio	0.04141	0.04854	0.04814	0.04547	0.03914
t-statistic	0.42401	0.65013	0.64417	0.55016	0.32835
p-value	0.67157	0.51563	0.51948	0.58223	0.74266

Notes to Table XI: The one-month treasury bill from Ibbotson Associates has been used as the risk-free rate. When computing the t-statistic all portfolios have been compared to the market portfolio. The portfolio used as a proxy for the market portfolio is the same as for the regressions. This a value-weighted portfolio consisting of all shares listed on NYSE, AMEX or NASDAQ. All numbers have been rounded to five decimals. * Significant at the 10 percent level, ** Significant at the 5 percent level, *** Significant at the 1 percent level.

The formula used for the t-test is: $t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$

As can be seen from Table X, the highest risk-adjusted returns among the best in class portfolios are achieved by the Best in Class Governance portfolio. However, all of these portfolios yield a lower Sharpe ratio compared to the Sector Exclusion portfolio. Looking at Table XI, it can be observed that the highest risk-adjusted return is demonstrated by the Worst in Class ESG portfolio. Overall the non-ESG portfolios exhibit higher Sharpe ratios than the ESG portfolios. What should be highlighted is that the two tables show that none of the reported Sharpe ratios are statistically significant. Since none of the Sharpe ratios are significant no further interpretation or analysis will be provided based on these.

Continuing with the Treynor ratios, these are presented in Tables XII and XIII. The ratios are computed based on data for the whole time period.

Table XII: Treynor Ratios ESG Portfolios

Variables	Sector Exclusion	Best in Class ESG	Best in Class Environmental	Best in Class Social	Best in Class Governance
Sharpe Ratio	0.05690	0.04741	0.04563	0.04294	0.05071
t-statistic	0.64110	0.32457	0.26122	0.16656	0.41793
p-value	0.52148	0.74551	0.79393	0.86772	0.67601

Notes to Table XII: The one-month treasury bill from Ibbotson Associates has been used as the risk-free rate. When computing the t-statistic all portfolios have been compared to the market portfolio. The portfolio used as a proxy for the market portfolio is the same as for the regressions. This a value-weighted portfolio consisting of all shares listed on NYSE, AMEX or NASDAQ. All numbers have been rounded to five decimals. * Significant at the 10 percent level, ** Significant at the 5 percent level, *** Significant at the 1 percent level.

The formula used for the t-test is: $t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$

Table XIII: Treynor Ratios non-ESG Portfolios

Variables	Conventional	Worst in Class ESG	Worst in Class Environmental	Worst in Class Social	Worst in Class Governance
Sharpe Ratio	0.05567	0.07002	0.06793	0.06694	0.05401
t-statistic	0.60188	1.06428	1.00577	0.96599	0.51705
p-value	0.54727	0.28724	0.31456	0.33408	0.60514

Notes to Table XIII: The one-month treasury bill from Ibbotson Associates has been used as the risk-free rate. When computing the t-statistic all portfolios have been compared to the market portfolio. The portfolio used as a proxy for the market portfolio is the same as for the regressions. This a value-weighted portfolio consisting of all shares listed on NYSE, AMEX or NASDAQ. All numbers have been rounded to five decimals. * Significant at the 10 percent level, ** Significant at the 5 percent level, *** Significant at the 1 percent level.

The formula used for the t-test is: $t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$

The Treynor ratios demonstrate similar results as the Sharpe ratios. Once again, it is the governance portfolio that has the highest ratio out of the best in class portfolios. Still, the Sector Exclusion portfolio has a higher Treynor ratio. This makes it the best-performing portfolio out of all the ESG portfolios. Just as for the Sharpe ratio, it is the Worst in Class ESG portfolio that has the highest Treynor ratio out of all portfolios. Overall the non-ESG portfolios seem to provide higher excess returns per unit of risk. However, no further interpretation or analysis will be presented since none of the portfolios has a ratio that is statistically significant.

5.3 Research Question II - Is there any difference in investment style between the portfolios? If so, how does the investment style differ?

Having examined the risk-adjusted returns, the focus now shifts to the investment styles of the different portfolios. Firstly, we will examine the beta of the portfolios. Both the best and worst in class portfolios based on governance have a relatively high beta. These portfolios have the highest beta in their respective table. Moreover, both betas are larger than one. Therefore, it is suggested that these portfolios are more volatile than the market portfolio. The results are thus in line with Ibikunle and Steffen (2015) who argue that the positive screening process might lead to the exclusion of traditional value stocks and, thus, a higher beta of the portfolio. In contrast, the remaining portfolios experience a beta smaller than one, which suggests that the volatility of these portfolios is lower than the market. All the best in class portfolios except the governance one have a lower beta than the Conventional portfolio. This is not in line with previous literature from (Climent and Soriano 2011; Ibikunle and Steffen 2015). According to their studies, the best in class portfolios should possess a higher beta than the Conventional portfolio. This is mainly due to the smaller investment universe. The results are therefore considered to be puzzling with regard to the theoretical reasoning provided in the literature. When the sample is split, the results for the first period align with the reasoning that a smaller investment universe should lead to a higher beta. In contrast, the second period is again not consistent with the examined literature.

Proceeding with the results for the SMB factors, the significant best in class portfolios show a negative SMB factor. This suggests a tilt toward large capitalization stocks. In contrast, the Sector Exclusion portfolio and the non-ESG portfolios display a positive SMB factor. The results for the whole time period are in line with the findings of Kaiser (2020), which show that larger firms are associated with higher ESG scores. As discussed in the literature review, the rationale provided is linked to increased shareholder pressure for larger companies. However, as previously discussed (Climent and Soriano 2011; Ibikunle and Steffen 2015) do not fully agree. The authors argue that positive screening processes have a tendency to exclude large capitalization stocks. If this would be seen in the results, the best in class portfolios should display positive SMB factors. Even if the two opposite effects are present in the results, the positive relationship suggested by Kaiser (2020) seems to dominate. Focusing on the subperiods, there appear to be no major time-varying effects.

Moving on from the SMB factor to the HML factor, the significant ESG-portfolios overall indicate a tilt toward value stocks. In contrast, the non-ESG portfolios that are significant have negative HML factors. This instead implies an overweight of growth stocks. The results for the best in class portfolios are not in line with the findings from (Ibikunle and Steffen 2015; Kaiser 2020). Those findings suggest that portfolios consisting of higher ESG-rated firms are associated with growth stocks. As mentioned in the literature review, the reasoning is based on the fact that growth companies tend to invest more in innovative environmental and clean tech. The former of the two papers also discuss how acquisitions can be a possible explanation for the somewhat time-varying effect on the growth stock bias. Furthermore, as discussed in relation to beta, a positive screening approach tends to exclude value companies (Ibikunle and Steffen 2015). However, our results point toward the opposite of these theoretical rationales. Thus, there seems to be a discrepancy between the empirical findings for the whole time period and the previous literature.

When dividing the time period, the results align themselves somewhat more with the literature. For the first period, all HML factors become negative for the best in class portfolios, implying a tilt toward growth stocks. The significant non-ESG portfolios are also consistent, with the literature exhibiting positive HML factors. During the second period, the signs for the ESG portfolios change. Half of the significant portfolios display a negative sign while the other half exhibit a positive sign. This implies both consistency and inconsistency with the previous literature discussed. In addition, the three significant non-ESG portfolios have negative HML factors. Once again, this suggests that the non-ESG portfolios are tilted toward growth stocks which are not in line with the examined literature.

The last of the four factors is the momentum factor. Out of the ESG portfolios, there are two that are significant. These are the best in class portfolios building upon ESG scores and environmental scores. Both these have a positive MOM factor which indicates a tilt toward positive momentum stocks. There is only one non-ESG portfolio with a significant MOM factor. This is the Worst in Class Governance portfolio, which has a negative sign. When comparing portfolios with high and low ESG scores (Ibikunle and Steffen 2015; Kaiser 2020), reach the conclusion that portfolios with positive momentum stocks are tilted away from high ESG scores. Our results are not consistent with those results. Therefore, the results do not coincide with the media coverage rationale provided by Kaiser (2020).

Our results do not show that negative momentum stocks try to compensate by improving their ESG performance. Neither does the result display that positive momentum stocks are less concerned about improving their ESG performance. Once again, our results imply that there are other aspects that affect the investment styles that are not explained by the previous literature.

As for the other factors, MOM is also derived for each of the two sub-periods. When looking at the first period, the results are partially in line with the literature. The significant ESG portfolios are tilted away from momentum stocks. However, the same holds for the non-ESG portfolios. Therefore the media coverage rationale from Kaiser (2020) does still not show in the results. In contrast, the results during the second period illustrate that the non-ESG portfolios are tilted toward momentum stocks. Similarly, the ESG portfolios also have a tilt toward momentum stocks. The results are, thus, partially consistent with the literature for the second period. Compared to the SMB factor, it is harder to explain the results for the HML and MOM factors with some sort of theoretical rationale. Furthermore, the results seem to be time-varying, which further complexifies the analysis. As will be more thoroughly discussed in section 6.1, we, therefore, encourage future research in the area.

5.4 Robustness

It can be observed from the results that all of the portfolios exhibit a positive alpha. Not all alphas are statistically significant, but nonetheless, no alpha is negative. Simplistically, the results suggest that all portfolios would at least not be worse performing than the market portfolio. An issue with the market portfolio is that it cannot be observed, and thus, proxies are used. This challenge was highlighted by Roll (1977). To make the results more robust, all the Carhart regressions are re-run using another market portfolio. Instead of using the proxy for the market portfolio from the Kenneth R. French database (see: section 3.3), the Conventional Portfolio will be used. To remind the readers, the Conventional Portfolio consists of all stocks in S&P 500 as of April 2023. Using S&P 500 as a proxy for the market portfolio is not unusual. For example, Albuquerque et al. (2020) use the S&P 500 as a proxy for the market portfolio. The results of these regressions are displayed in Table D1-D6 in Appendix D.

All the alphas for all the best in class portfolios become negative when changing the market portfolio. Furthermore, the Sector Exclusion portfolio is the only ESG portfolio that has a positive alpha. An important note is that there is only one alpha that is statistically significant. This is for the best in class social portfolio. Moreover, it is just significant on a ten percent level. The alphas for the worst in class portfolios are all positive except for one portfolio. This is for the Worst in Class Governance portfolio. In addition, two of the four portfolios are significant. It is important to recognize that the significance level is just at ten percent. By just focusing on the signs, the previous results and analysis presented seem to be consistent even when changing the market portfolio. However, due to the statistical significance, no deeper analysis or interpretation will be provided.

The main results for the investment styles of the portfolios seem to be robust even when changing the market portfolio. The betas are still overall relatively low for the best in class portfolios which is puzzling. Furthermore, these portfolios are also tilted toward large capitalization stocks, while the worst in class portfolios are tilted toward small capitalization stocks. This holds except for the second subperiod. The HML and MOM factors, in general, also have the same signs and show some time-varying changes.

6. Conclusion

The first research question that has been examined is how the risk-adjusted returns differ between the constructed portfolios. By using data with a higher frequency and examining ten different portfolios, we have contributed with an enhanced perspective on how various ESG approaches affect risk-adjusted returns. It can be concluded that the results point towards that the non-ESG portfolios, in general, perform better than the ESG portfolios. It is possible that investors only have one goal, which is to maximize risk-adjusted returns. These investors should, according to the results, be careful about implementing ESG into their portfolios. The results seem to be persistent both with regard to different models as well as time periods. Different theoretical rationales have been provided in the analysis, which tries to explain these findings.

Another insight is that the findings suggest that a negative screening approach is superior to a positive screening approach.

Therefore, investors that wish to implement ESG into their portfolios should mainly consider excluding certain industries instead of choosing companies with the best scores. Again, these findings are consistent both when applying different models and time periods. A detailed analysis has been provided with regard to the different ESG pillars. The findings have pointed toward that investing in companies with excellent governance can be a relatively good ESG approach. A remark to consider is that this procedure showed to be time-varying. Potential explanations have been provided to understand the results. However, independent of the theoretical explanation behind the results, they still suggest a takeaway. If investors desire to follow a non-ESG approach, investing in companies with bad governance is not optimal. Instead, investors should focus on a non-ESG approach built upon combined ESG scores or one of the other pillars.

The implications from the key results seem to be unchanged when changing the market portfolio if just looking at the signs. Also, when focusing on the magnitude of the financial ratios, there appears to be no difference. However, due to the statistical significance, we are careful about drawing any conclusions based on those results. We also want to remind the readers that all results build upon realized returns. Therefore, one should distinguish these from the expected returns. Even if, for example, the results point toward that pursuing a non-ESG approach has been a good idea for investors, extrapolating this into the future can be risky. A detailed discussion regarding this topic was also provided in the literature review.

The second research question this thesis examines is if there is any difference in investment style between the portfolios and, if so, how these differ. We expand the literature by examining investment styles with regard to multiple strategies. It is revealed from the results that the styles of the constructed portfolios differ. By interpreting the factors from the Carhart model for the whole time period, it can be seen that overall the best in class portfolios exhibit low betas. This hold but for the whole period as well as for the second subperiod. Looking at the other factors, ESG portfolios are generally tilted toward large capitalization and value stocks. Moreover, these portfolios also have a larger exposure toward positive momentum stocks. The results also suggest that the non-ESG portfolios have a tilt toward small capitalization and growth stocks. The result also suggests exposure away from positive momentum stocks except for the second subperiod. Some time-varying effects are found and the most noticeable are connected to the HML and MOM factors.

As previously discussed, the Morningstar Style Box has for a long time acted as a supporting tool for categorizing and evaluating mutual funds' investment styles. In a similar way, this thesis can hopefully be a guiding principle for investors with regard to the construction of ESG portfolios. Investors that wish to pursue some sort of ESG investing approach can potentially benefit from this thesis by being more aware of in what way particular strategies can imply potential investment styles. Furthermore, the thesis uncovers some of the uncertainty associated with the time-varying effects. Hopefully, this can also help investors in their decision process regarding if and how to implement ESG into their investment strategy.

6.1 Limitations and Future Research

The fundamental problems with ESG scores are an aspect that has been discussed previously. Due to the subjective nature, it is possible that if this study is repeated using another data provider, the results can differ. Furthermore, the addressed issue regarding potential look-ahead bias and survivorship bias can also be considered a limitation. Based on the limitations of the thesis, it would be interesting to see the results from a replicated study using another data provider. Moreover, it could also be the case that the data provider certainly keeps the issued scores and doesn't change them retroactively. This would help to control for look-ahead bias. In this thesis, no rational explanation could be provided for the time-varying aspects of the portfolio building upon the companies with the best governance scores. We, therefore, also encourage future research focusing on the specific pillars and any eventual time-varying aspects. Furthermore, some of the results linked to the styles can be considered puzzling. That the best in class portfolios exhibit higher betas could not be logically explained by the examined theory and literature. The same goes for the momentum factors during both the whole and first time period.

Our impression is that, in general, there is less research in the area of investment styles linked to ESG investing. Especially there are fewer theoretical papers that try to explain the empirical findings in the literature. As a result, there could be opportunities for further research with regard to that. Lastly, the impact of our discoveries remains constrained by the data at hand and is solely valid for our particular timeframe and the geographic area we have examined. Expanding the study to different time periods and other geographic areas could be appealing for future research.

References

Albuquerque, R., Koskinen, Y., Yang, S. and Zhang, C. (2020). Resiliency of Environmental and Social Stocks: An Analysis of the Exogenous COVID-19 Market Crash. *The Review of Corporate Finance Studies*, 9(3), pp. 593–621. doi:https://doi.org/10.1093/rcfs/cfaa011.

Amon, J., Rammerstorfer, M. and Weinmayer, K. (2021). Passive ESG Portfolio Management—The Benchmark Strategy for Socially Responsible Investors. *Sustainability*, 13(16), p. 9388. doi:https://doi.org/10.3390/su13169388.

Auer, B.R. (2014). Do Socially Responsible Investment Policies Add or Destroy European Stock Portfolio Value? *Journal of Business Ethics*, 135(2), pp. 381–397. doi:https://doi.org/10.1007/s10551-014-2454-7.

Banz, R.W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, 9(1), pp. 3–18. doi:https://doi.org/10.1016/0304-405x(81)90018-0.

Barnett, M.L. and Salomon, R.M. (2006). Beyond dichotomy: the curvilinear relationship between social responsibility and financial performance. *Strategic Management Journal*, 27(11), pp. 1101–1122. doi:https://doi.org/10.1002/smj.557.

Bauer, R., Koedijk, K. and Otten, R. (2005). International evidence on ethical mutual fund performance and investment style. *Journal of Banking & Finance*, 29(7), pp. 1751–1767. doi:https://doi.org/10.1016/j.jbankfin.2004.06.035.

Bender, J., Bridges, T.A., He, C., Lester, A. and Sun, X. (2017). A Blueprint for Integrating ESG into Equity Portfolios. *The Journal of Investment Management*, 16(1), pp. 44-58. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3080381.

Bennani, L., Le Guenedal, T., Lepetit, F., Ly, L., Mortier, V., Roncalli, T. and Sekine, T. (2018). How ESG Investing Has Impacted the Asset Pricing in the Equity Market. *SSRN Electronic Journal*. doi:https://doi.org/10.2139/ssrn.3316862.

Berg, F., Kölbel, J. and Rigobon, R. (2019). Aggregate Confusion: The Divergence of ESG Ratings. *SSRN Electronic Journal*. doi:https://doi.org/10.2139/ssrn.3438533.

Blázquez-García, A., Conde, A., Mori, U. and Lozano, J.A. (2021). A Review on Outlier/Anomaly Detection in Time Series Data. *ACM Computing Surveys*, 54(3), pp. 1–33. doi:https://doi.org/10.1145/3444690.

Bodie, Z., Kane, A. and Marcus, A.J. (2020) *Investments*. 12th ed., McGraw Hill Education.

Boffo, R., and R. Patalano (2020), "ESG Investing: Practices, Progress and Challenges", OECD Paris, www.oecd.org/finance/ESG-Investing-Practices-Progress-and-Challenges.pdf

Brown, S.J., Goetzmann, W., Ibbotson, R.G. and Ross, S.A. (1992). Survivorship Bias in Performance Studies. *Review of Financial Studies*, 5(4), pp. 553–580. doi:https://doi.org/10.1093/rfs/5.4.553.

Brammer, S.J., Brooks, C. and Pavelin, S. (2006). Corporate Social Performance and Stock Returns: UK Evidence from Disaggregate Measures. *Financial Management,* 35(3), pp. 97-116. doi:https://doi.org/10.2139/ssrn.739587.

Capelle-Blancard, G. and Monjon, S. (2012). The Performance of Socially Responsible Funds: Does the Screening Process Matter? *European Financial Management*, 20(3), pp.494–520. doi:https://doi.org/10.1111/j.1468-036x.2012.00643.x.

Carhart, M.M. (1997). On Persistence in Mutual Fund Performance. *The Journal of Finance*, 52(1), p. 57. doi:https://doi.org/10.2307/2329556.

Christie, S 2005, Is the Sharpe Ratio Useful in Asset Allocation?, *MAFC Research Papers*, pp. 1-48. doi:https://doi.org/10.2139/ssrn.720801.

Climent, F. and Soriano, P. (2011). Green and Good? The Investment Performance of US Environmental Mutual Funds. *Journal of Business Ethics*, 103(2), pp. 275–287. doi:https://doi.org/10.1007/s10551-011-0865-2.

Daniel, G., Sornette, D. and Woehrmann, P. (2009). Look-Ahead Benchmark Bias in Portfolio Performance Evaluation. *The Journal of Portfolio Management*, 36(1), pp. 121–130. doi:https://doi.org/10.3905/jpm.2009.36.1.121.

Derwall, J., Guenster, N., Bauer, R. and Koedijk, K. (2005). The Eco-Efficiency Premium Puzzle. *Financial Analysts Journal*, 61(2), pp. 51–63. doi:https://doi.org/10.2469/faj.v61.n2.2716.

Dorfleitner, G., Kreuzer, C. and Sparrer, C. (2020). ESG controversies and controversial ESG: about silent saints and small sinners. *Journal of Asset Management*, 21(5), pp. 393–412. doi:https://doi.org/10.1057/s41260-020-00178-x.

Drei, A., Le Guenedal, T., Lepetit, F., Mortier, V., Roncalli, T. and Sekine, T. (2019). ESG Investing in Recent Years: New Insights from Old Challenges. *SSRN Electronic Journal*. doi:https://doi.org/10.2139/ssrn.3683469.

Ehlers, T., Elsenhuber, U., Jegarasasingam, K. and Jondeau, E. (2023). Deconstructing ESG Scores: How to Invest with your own Criteria? *IMF Working Papers*, 2023(057), p. 1. doi:https://doi.org/10.5089/9798400235283.001.

Elton, E.J. (1999). Presidential Address: Expected Return, Realized Return, and Asset Pricing Tests. *The Journal of Finance*, 54(4), pp. 1199–1220. doi:https://doi.org/10.1111/0022-1082.00144.

Fama, E.F. and French, K.R. (1992). The Cross-Section of Expected Stock Returns. *The Journal of Finance*, 47(2), pp. 427–465. doi:https://doi.org/10.2307/2329112.

Fama, E.F. and French, K.R. (2002). The Equity Premium. *The Journal of Finance*, 57(2), pp. 637–659. doi:https://doi.org/10.1111/1540-6261.00437.

Fama, E.F. and French, K.R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1), pp. 1–22. doi:https://doi.org/10.1016/j.jfineco.2014.10.010.

Friede, G., Busch, T. and Bassen, A. (2015). ESG and financial performance: aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment*, 5(4), pp. 210–233. doi: https://doi.org/10.1080/20430795.2015.1118917

French, K.R. (2023). *Kenneth R. French - Data Library*. Available at: https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Grim, D.M. and Berkowitz, D.B. (2020). ESG, SRI, and Impact Investing: A Primer for Decision-Making. *The Journal of Impact and ESG Investing*, 1(1), pp. 47–65. doi:https://doi.org/10.3905/jesg.2020.1.1.047.

Henriksson, R., Livnat, J., Pfeifer, P. and Stumpp, M. (2019). Integrating ESG in Portfolio Construction. *The Journal of Portfolio Management*, 45(4), pp. 67–81. doi:https://doi.org/10.3905/jpm.2019.45.4.067.

Hillman, A.J. and Keim, G.D. (2001). Shareholder value, stakeholder management, and social issues: what's the bottom line? *Strategic Management Journal*, 22(2), pp. 125–139. doi: https://doi.org/10.1002/1097-0266(200101)22:2%3C125::AID-SMJ150%3E3.0.CO;2-H

Hoepner, A.G.F., Oikonomou, I., Sautner, Z., Starks, L.T. and Zhou, X. (2016). ESG Shareholder Engagement and Downside Risk. *SSRN Electronic Journal*. doi:https://doi.org/10.2139/ssrn.2874252.

Ibikunle, G. and Steffen, T. (2015). European Green Mutual Fund Performance: A Comparative Analysis with their Conventional and Black Peers. *Journal of Business Ethics*, 145(2), pp. 337–355. doi:https://doi.org/10.1007/s10551-015-2850-7.

Jegadeesh, N. and Titman, S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *The Journal of Finance*, 48(1), pp. 65–91. doi:https://doi.org/10.2307/2328882.

Jennings, K. (2008). *Covered Short Selling by UCITS Funds - The Debate Continues*. Dillon Eustace. https://www.dilloneustace.com/download/1/Covered%20Short%20Selling%20by%20UCITS.pdf.

Jones, C.M. and Lamont, O.A. (2002). Short-sale constraints and stock returns. *Journal of Financial Economics*, 66(2-3), pp. 207–239. doi:https://doi.org/10.1016/s0304-405x(02)00224-6.

Kaiser, L. (2020). ESG integration: value, growth and momentum. *Journal of Asset Management*, 21(3), pp. 32–51. doi:https://doi.org/10.1057/s41260-019-00148-y.

Kempf, A. and Osthoff, P. (2007). The Effect of Socially Responsible Investing on Portfolio Performance. *European Financial Management*, 13(5), pp. 908–922. doi:https://doi.org/10.1111/j.1468-036x.2007.00402.x.

Kinnel, R. (2002). The New Morningstar Style Box. Morningstar Inc*,* June 5. https://www.morningstar.com/articles/44956/the-new-morningstar-style-box.

Kumar, A. (2009). Dynamic Style Preferences of Individual Investors and Stock Returns. *Journal of Financial and Quantitative Analysis*, 44(3), pp. 607–640. doi:https://doi.org/10.1017/s0022109009990020.

Ledoit, O. and Wolf, M. (2008). Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance*, 15(5), pp. 850–859. doi:https://doi.org/10.1016/j.jempfin.2008.03.002.

Lintner, J. (1965). The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *The Review of Economics and Statistics*, 47(1), pp. 13–37. doi:https://doi.org/10.2307/1924119.

Lioui, A. and Tarelli, A. (2022). Chasing the ESG factor. *Journal of Banking & Finance*, 139, p.106498. doi:https://doi.org/10.1016/j.jbankfin.2022.106498.

Mănescu, C. (2011). Stock returns in relation to environmental, social and governance performance: Mispricing or compensation for risk? *Sustainable Development*, 19(2), pp. 95–118. doi:https://doi.org/10.1002/sd.510.

Mardini, G.H. (2022). ESG Factors and Corporate Financial Performance. *International Journal of Managerial and Financial Accounting*, 14(3), pp 247-264. doi: https://doi.org/10.1504/ijmfa.2022.123895.

Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, 7(1), pp. 77–91. doi:https://doi.org/10.2307/2975974.

Mossin, J. (1966). Equilibrium in a Capital Asset Market. *Econometrica*, 34(4), p. 768. doi:https://doi.org/10.2307/1910098.

MSCI (2023). MSCI ACWI Index (USD) [Fact sheet]. https://www.msci.com/documents/10199/8d97d244-4685-4200-a24c-3e2942e3adeb

Mushtaq, R. (2011). Augmented Dickey Fuller Test. *SSRN Electronic Journal*. doi:https://doi.org/10.2139/ssrn.1911068.

Nagy, Z., Kassam, A. and Lee, L.-E. (2016). Can ESG Add Alpha? An Analysis of ESG Tilt and Momentum Strategies. *The Journal of Investing*, 25(2), pp. 113–124. doi:https://doi.org/10.3905/joi.2016.25.2.113.

Opdyke, J.D. (2007). Comparing Sharpe ratios: So where are the p-values? *Journal of Asset Management*, 8(5), pp. 308–336. doi:https://doi.org/10.1057/palgrave.jam.2250084.

Pástor, Ľ., Stambaugh, R.F. and Taylor, L.A. (2020). Sustainable Investing in Equilibrium. *Journal of Financial Economics*, 142(2), pp. 550-571. doi:https://doi.org/10.1016/j.jfineco.2020.12.011.

Pástor, Ľ., Stambaugh, R.F. and Taylor, L.A. (2022). Dissecting green returns. *Journal of Financial Economics*, 146(2), pp. 403–424. doi:https://doi.org/10.1016/j.jfineco.2022.07.007.

Pedersen, L.H., Fitzgibbons, S. and Pomorski, L. (2020). Responsible investing: The ESG-efficient frontier. *Journal of Financial Economics*, 142(2), pp. 572–597. doi:https://doi.org/10.1016/j.jfineco.2020.11.001.

Refinitiv (2022). Environmental, Social and Governance Scores From Refinitiv [Fact sheet]. https://www.refinitiv.com/content/dam/marketing/en_us/documents/methodology/refinitiv-esg-scores-methodology.pdf.

Renneboog, L., Ter Horst, J. and Zhang, C. (2008). The price of ethics and stakeholder governance: The performance of socially responsible mutual funds. *Journal of Corporate Finance*, 14(3), pp. 302–322. doi:https://doi.org/10.1016/j.jcorpfin.2008.03.009.

Revelli, C. and Viviani, J.-L. (2014) Financial Performance of Socially Responsible Investing (SRI): What Have We Learned? A Meta-Analysis, 24(2), pp. 158-186. *Business Ethics.* doi:https://doi.org/10.1111/beer.12076

Roll, R. (1977). A critique of the asset pricing theory's tests Part I: On past and potential testability of the theory. *Journal of Financial Economics*, 4(2), pp. 129–176. doi:https://doi.org/10.1016/0304-405x(77)90009-5.

Schröder, M. (2007). Is there a Difference? The Performance Characteristics of SRI Equity Indices. *Journal of Business Finance & Accounting*, 34(1-2), pp. 331–348. doi:https://doi.org/10.1111/j.1468-5957.2006.00647.x.

S&P Dow Jones Indices (2023a). Index Mathematics Methodology [Fact Sheet]. https://www.spglobal.com/spdji/en/documents/methodologies/methodology-index-math.pdf

S&P Dow Jones Indices (2023b). S&P ESG Leaders Indices Methodology [Fact Sheet]. https://www.spglobal.com/spdji/en/documents/methodologies/methodology-sp-esg-leaders-indices.pdf

Sharpe, W.F. (1964). Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk. *The Journal of Finance*, 19(3), pp. 425–442. doi:https://doi.org/10.1111/j.1540-6261.1964.tb02865.x.

Statman, M. and Glushkov, D. (2009). The Wages of Social Responsibility. *Financial Analysts Journal*, 65(4), pp. 33–46. doi:https://doi.org/10.2469/faj.v65.n4.5.

Stock, J.H. and Watson, M.W. (2019). *Introduction to Econometrics*. 4th ed., Pearson.

Uyanto, S.S. (2020). Power Comparisons of Five Most Commonly Used Autocorrelation Tests. *Pakistan Journal of Statistics and Operation Research*, 16(1), pp. 119–130. doi:https://doi.org/10.18187/pjsor.v16i1.2691.

Wright, J.A., Yam, S.C.P. and Yung, S.P. (2014). A test for the equality of multiple Sharpe ratios. *The Journal of Risk*, 16(4), pp. 3–21. doi:https://doi.org/10.21314/jor.2014.289.

Wu, J. (2022). *ESG Outlook 2022: The future of ESG investing*. J.P. Morgan Asset Management. https://am.jpmorgan.com/se/en/asset-management/liq/investment-themes/sustainable-investing/future-of-esg-investing/.

Ziegler, A., Schröder, M. and Rennings, K. (2007). The effect of environmental and social performance on the stock performance of European corporations. *Environmental and Resource Economics*, 37(4), pp. 661–680. doi:10.1007/s10640-007-9082-y.

Appendix

Appendix A:

Table A1 – ESG Terminology

Terminology	Description
Socially Responsible Investing (SRI)	Socially responsible investing (SRI) is closely related to ESG investing. Just as ESG investing, SRI is a broad term. The essence of SRI is different portfolio screens built on ethical, social, corporate governance or environmental criteria.
Sin Stocks	Sin Stocks are defined as companies that operate in certain industries. Examples of these industries are: tobacco, alcohol, arms or gambling industry.
“Green” Firms	“Green” firms are defined as firms that create positive externalities for society.
“Brown” Firms	“Brown” firms are defined as firms that create negative externalities for society.
“Black” Peers	“Black” peers are defined as mutual funds that invest in certain companies. These companies are operating in business areas that involve exploitation and depletion of nature. For example, it can be companies involved in the oil, gas and/or coal industry.

Notes to Table A1: Sources: (Dorfleitner, Kreuzer and Sparrer 2020; Grim and Berkowitz 2020; Ibikunle and Steffen 2015; Pástor, Stambaugh and Taylor 2020; Renneboog, Ter Horst and Zhang 2008)

Table A2 – Refinitiv Eikon Score Structure

Category	Subcategory	ESG metrics	Percentage (%) weights
Environmental	Resource use	20	11
	Emissions	28	15
	Innovation	20	11
Social	Workforce	30	16
	Human rights	8	4
	Community	14	8
	Product responsibility	10	5
Governance	Management	35	19
	Shareholders	12	6
	CSR strategy	9	5
Total		**186**	**100**

Notes to Table A2: Sources: (Refinitiv Eikon 2022)

Table A3 – Excluded Industries

Product Involvement	Category of Involvement and Description
Tobacco products	**Production:** The company manufactures tobacco products.
	Related Products/Services: The company supplies tobacco-related products/services.
	Retail: The company derives revenues from the distribution and/or retail sale of tobacco products.
Oil Sands	**Extraction:** The company extracts oil sands.
Shale Energy	**Extraction:** The company is involved in shale energy exploration and/or production.
Arctic Oil & Gas Exploration	**Extraction:** The company is involved in oil and gas exploration in Arctic regions.
Small Arms	**Civilian Customers (Assault Weapons):** The company manufactures and sells assault weapons to civilian customers.
	Civilian customers (Non-assault weapons): The company manufactures and sells small arms (Non-assault weapons) to civilian customers.
	Military/law enforcement customers: The company manufactures and sells small arms to military/law enforcement customers.
	Key Components: The company manufactures and sells key components of small arms.
	Retail/distribution (Assault weapons): The company is involved in the retail and/or distribution of assault weapons.
	Retail/distribution (Non-assault weapons): The company is involved in the retail and/or distribution of small arms (Non-assault weapons).
Alcoholic Beverages	**Production:** The company manufactures alcoholic beverages.
	Retail (≥10% total revenues): The company derives revenues from the distribution and/or retail sale of alcoholic beverages.
	Related Products/Services: The company is a supplier of alcohol-related products/services to alcoholic beverage manufacturers.
Gambling	**Operations:** The company owns and/or operates a gambling establishment.
	Specialized Equipment: The company manufactures specialized equipment used exclusively for gambling.
	Supporting Products/Services: The company provides supporting products/services to gambling operations.
Nuclear Power	**Production:** The company produces nuclear power.
	Distribution: The company distributes electricity generated from nuclear power.
	Supporting Products/Services: The company provides products/services that support the nuclear power industry.

Notes to Table A2: Sources: (S&P Dow Jones Indices 2023b)

Appendix B:

Table B1 – Augmented Dickey Fuller Test for the ESG portfolios

Portfolio	Sector Exclusion	Best in Class ESG	Best in Class Environmental	Best in Class Social	Best in Class Governance
Test Statistic	-65.936***	-67.147***	-67.797***	-68.499***	-67.102***

Notes to table B1: $H_0 = \ Random\ walk\ without\ drift$

* Significant at a 10 percent level, ** Significant at a 5 percent level, *** Significant at a 1 percent level.

Table B2 – Augmented Dickey Fuller Test for the non-ESG portfolios

Portfolio	Conventional	Worst in Class ESG	Worst in Class Environmental	Worst in Class Social	Worst in Class Governance
Test Statistic	-66.247***	-62.199***	-62.665***	-62.332***	-65.475***

Notes to table B2: $H_0 = \ Random\ walk\ without\ drift$

* Significant at a 10 percent level, ** Significant at a 5 percent level, *** Significant at a 1 percent level.

Table B3 – Breusch-Godfrey Test for the ESG portfolios

Portfolio	Sector Exclusion	Best in Class ESG	Best in Class Environmental	Best in Class Social	Best in Class Governance
P-value	0.8402	0.2104	0.0563*	0.0723*	0.0786*

Notes to table B3: $H_0 = \ No\ serial\ Correlation$

* Significant at a 10 percent level, ** Significant at a 5 percent level, *** Significant at a 1 percent level.

Table B4 – Breusch-Godfrey Test for the non-ESG portfolios

Portfolio	Conventional	Worst in Class ESG	Worst in Class Environmental	Worst in Class Social	Worst in Class Governance
P-value	0.2332	0.6335	0.6622	0.7616	0.000***

Notes to table B4: $H_0 = \ No\ serial\ Correlation$

* Significant at a 10 percent level, ** Significant at a 5 percent level, *** Significant at a 1 percent level.

Table B5 – Breusch-Pagan/Cook-Weisberg Test for the ESG portfolios

Portfolio	Sector Exclusion	Best in Class ESG	Best in Class Environmental	Best in Class Social	Best in Class Governance
Chi2(4)	491.65***	75.05***	223.11***	95.61***	261.99***
P-value	0.0000	0.0000	0.0000	0.0000	0.0000

Notes to table B5: $H_0 = Constant\ Variance$. The test has been conducted on all the independent variables in the Carhart model.

* Significant at a 10 percent level, ** Significant at a 5 percent level, *** Significant at a 1 percent level.

Table B6 – Breusch-Pagan/Cook-Weisberg Test for the ESG portfolios

Portfolio	Conventional	Worst in Class ESG	Worst in Class Environmental	Worst in Class Social	Worst in Class Governance
Chi2(4)	296.55***	848.85***	579.51***	725.85***	1552.40***
P-value	0.0000	0.0000	0.0000	0.0000	0.0000

Notes to table B6: $H_0 = Constant\ Variance$. The test has been conducted on all the independent variables in the Carhart model.

* Significant at a 10 percent level, ** Significant at a 5 percent level, *** Significant at a 1 percent level.

Appendix C:

Table C1 – Descriptive Statistics Subperiods ESG Portfolios

Variables	Sector Exclusion		Best in Class ESG		Best in Class Environmental		Best in Class Social		Best in Class Governance	
	2007-2014	2015-2022	2007-2014	2015-2022	2007-2014	2015-2022	2007-2014	2015-2022	2007-2014	2015-2022
Mean	0.05476	0.05562	0.04121	0.04704	0.04381	0.04110	0.03702	0.04490	0.04153	0.06120
Median	0.11543	0.10156	0.09915	0.07991	0.09574	0.06417	0.09333	0.07310	0.09618	0.10053
St. Dev	1.40059	1.21121	1.31474	1.16490	1.36983	1.11756	1.36309	1.13803	1.53723	1.20198
Min	-9.16086	-12.52766	-8.52054	-11.67653	-8.97281	-11.31580	-8.93320	-10.43327	-9.93693	-12.18762
Max	9.19384	10.28456	10.22028	9.84840	10.83815	8.26030	11.09206	8.24708	11.65520	9.91572
Skewness	-0.13466	-0.61414	-0.13973	-0.42165	-0.10555	-0.66088	-0.08318	-0.48799	-0.12399	-0.55872
Kurtosis	9.60135	16.96567	11.20962	17.91060	11.85212	17.12651	12.63035	13.90945	11.35831	15.42288

Notes to table C1: The table displays descriptive statistics for the excess returns of each ESG portfolio. All numbers have been rounded to five decimals. The number of observations for each subperiod is 2014.

Table C2 – Descriptive Statistics Subperiods non-ESG Portfolios

Variables	Conventional		Worst in Class ESG		Worst in Class Environmental		Worst in Class Social		Worst in Class Governance	
	2007-2014	2015-2022	2007-2014	2015-2022	2007-2014	2015-2022	2007-2014	2015-2022	2007-2014	2015-2022
Mean	0.05331	0.05405	0.06597	0.06733	0.06421	0.06506	0.06222	0.06145	0.06031	0.05198
Median	0.11655	0.09254	0.10490	0.10392	0.09563	0.11412	0.10613	0.09774	0.13982	0.09147
St. Dev	1.38936	1.19669	1.42592	1.31868	1.38998	1.29424	1.40934	1.30928	1.59826	1.24989
Min	-9.10885	-12.54160	-9.22080	-13.67849	-9.51361	-14.67725	-9.06781	-14.36198	-10.02989	-12.13761
Max	9.37951	10.31037	13.34239	11.23332	13.10246	11.59299	13.36991	11.83516	13.76405	9.82920
Skewness	-0.16151	-0.64379	0.21825	-0.65235	0.15970	-0.68541	0.21266	-0.66018	0.15585	-0.50408
Kurtosis	9.89008	17.84201	11.78075	16.37218	11.75171	19.99905	11.90916	19.14421	12.16127	13.75931

Notes to table C2: The table displays descriptive statistics for the excess returns of each ESG portfolio. All numbers have been rounded to five decimals. The number of observations for each subperiod is 2014.

Table C3: ESG Portfolios CAPM

Variables	Sector Exclusion	Best in Class ESG	Best in Class Environmental	Best in Class Social	Best in Class Governance
Alpha	0.01796***	0.00898**	0.00692*	0.00524	0.01259**
	(0.00428)	(0.00449)	(0.00415)	(0.00406)	(0.00564)
Mkt-RF	0.97539***	0.92076***	0.93083***	0.93592***	1.01487***
	(0.00844)	(0.00541)	(0.00506)	(0.00534)	(0.00730)
R^2	0.9564	0.9470	0.9555	0.9575	0.9323

Notes to table C3: Robust standard errors are displayed in parentheses. All coefficients and robust standard errors are rounded to five decimals. All R^2 are rounded to four decimals. The number of observations is 4028 for all the portfolios.

* Significant at a 10 percent level, ** Significant at a 5 percent level, *** Significant at a 1 percent level.

Table C4: Non-ESG Portfolios CAPM

Variables	Conventional	Worst in Class ESG	Worst in Class Environmental	Worst in Class Social	Worst in Class Governance
Alpha	0.01676***	0.03026***	0.02972***	0.02596***	0.01643**
	(0.00411)	(0.00881)	(0.00898)	(0.00893)	(0.00682)
Mkt-RF	0.96722***	0.95339***	0.92516***	0.93990***	1.04076
	(0.00720)	(0.01860)	(0.01913)	(0.01923)	(0.01020)
R^2	0.9590	0.8305	0.8175	0.8228	0.9071

Notes to table C4: Robust standard errors are displayed in parentheses. All coefficients and robust standard errors are rounded to five decimals. All R^2 are rounded to four decimals. The number of observations is 4028 for all the portfolios.

* Significant at a 10 percent level, ** Significant at a 5 percent level, *** Significant at a 1 percent level.

Table C5: ESG Portfolios 2007-2014 CAPM

Variables	Sector Exclusion	Best in Class ESG	Best in Class Environmental	Best in Class Social	Best in Class Governance
	2007-2014	2007-2014	2007-2014	2007-2014	2007-2014
Alpha	0.02045***	0.00875	0.00987*	0.00325	0.00392
	(0.00693)	(0.00547)	(0.00509)	(0.00509)	(0.00815)
Mkr-RF	0.97229***	0.91986***	0.96194***	0.95703***	1.06306***
	(0.01313)	(0.00621)	(0.00569)	(0.00642)	(0.00951)
R^2	0.9499	0.9649	0.9720	0.9716	0.9431

Notes to table C5: Robust standard errors are displayed in parentheses. All coefficients and robust standard errors are rounded to five decimals. All R^2 are rounded to four decimals. The number of observations is 2014 for all the portfolios.

* Significant at a 10 percent level, ** Significant at a 5 percent level, *** Significant at a 1 percent level.

Table C6: Non-ESG Portfolios 2007-2014 CAPM

Variables	Conventional	Worst in Class ESG	Worst in Class Environmental	Worst in Class Social	Worst in Class Governance
	2007-2014	2007-2014	2007-2014	2007-2014	2007-2014
Alpha	0.01924***	0.03433**	0.03407**	0.03083**	0.02220**
	(0.00676)	(0.01476)	(0.01437)	(0.01440)	(0.01104)
Mkr-RF	0.96550***	0.89677***	0.87540***	0.88954***	1.08066***
	(0.01064)	(0.02783)	(0.02713)	(0.02748)	(0.01550)
R^2	0.9519	0.7796	0.7813	0.7852	0.9016

Notes to table C6: Robust standard errors are displayed in parentheses. All coefficients and robust standard errors are rounded to five decimals. All R^2 are rounded to four decimals. The number of observations is 2014 for all the portfolios.

* Significant at a 10 percent level, ** Significant at a 5 percent level, *** Significant at a 1 percent level.

Table C7: ESG Portfolios 2015-2022 CAPM

Variables	Sector Exclusion	Best in Class ESG	Best in Class Environmental	Best in Class Social	Best in Class Governance
	2015-2022	2015-2022	2015-2022	2015-2022	2015-2022
Alpha	0.01540***	0.00919	0.00457	0.00764	0.02203***
	(0.00501)	(0.00713)	(0.00640)	(0.00626)	(0.00745)
Mkr-RF	0.97954***	0.92195***	0.88929***	0.90772***	0.95049***
	(0.00870)	(0.00949)	(0.00849)	(0.00725)	(0.00908)
R^2	0.9651	0.9243	0.9343	0.9387	0.9222

Notes to table C7: Robust standard errors are displayed in parentheses. All coefficients and robust standard errors are rounded to five decimals. All R^2 are rounded to four decimals. The number of observations is 2014 for all the portfolios.

* Significant at a 10 percent level, ** Significant at a 5 percent level, *** Significant at a 1 percent level.

Table C8: Non-ESG Portfolios 2015-2022 CAPM

Variables	Conventional	Worst in Class ESG	Worst in Class Environmental	Worst in Class Social	Worst in Class Governance
	2015-2022	2015-2022	2015-2022	2015-2022	2015-2022
Alpha	0.01425***	0.02508***	0.02435**	0.02009*	0.01144
	(0.00470)	(0.00923)	(0.01048)	(0.01025)	(0.00782)
Mkr-RF	0.96953**	1.02903***	0.99160***	1.00720***	0.98749***
	(0.00885)	(0.01676)	(0.02080)	(0.02079)	(0.00949)
R^2	0.9685	0.8985	0.8662	0.8732	0.9210

Notes to table C8: Robust standard errors are displayed in parentheses. All coefficients and robust standard errors are rounded to five decimals. All R^2 are rounded to four decimals. The number of observations is 2014 for all the portfolios.

* Significant at a 10 percent level, ** Significant at a 5 percent level, *** Significant at a 1 percent level.

Appendix D:

Table D1: Changed Market Portfolio (S&P 500) ESG Portfolios Carhart

Variables	Sector Exclusion	Best in Class ESG	Best in Class Environmental	Best in Class Social	Best in Class Governance
Alpha	0.00146	-0.00592	-0.00726	-0.00991*	-0.00370
	(0.00218)	(0.00491)	(0.00502)	(0.00508)	(0.00548)
Mkt-RF	1.00083***	0.93546***	0.93508***	0.94939***	1.02907***
	(0.00248)	(0.00793)	(0.00875)	(0.00956)	(0.00861)
SMB	0.01860***	-0.11602***	-0.16948***	-0.18336***	-0.05325***
	(0.00402)	(0.01626)	(0.01487)	(0.01511)	(0.01463)
HML	-0.01089***	0.02146**	0.06349***	-0.05457***	0.01793*
	(0.00282)	(0.00933)	(0.00984)	(0.01005)	(0.01026)
MOM	-0.01184***	0.01188*	0.00386	-0.01997***	-0.01129
	(0.00245)	(0.00673)	(0.00832)	(0.00753)	(0.00845)
R^2	0.9887	0.9361	0.9343	0.9331	0.9367

Notes to table D1: Robust standard errors are displayed in parentheses. All coefficients and robust standard errors are rounded to five decimals. All R^2 are rounded to four decimals. The number of observations is 4028 for all the portfolios.

* Significant at a 10 percent level, ** Significant at a 5 percent level, *** Significant at a 1 percent level.

Table D2: Changed Market Portfolio (S&P 500) Non-ESG Portfolios Carhart

Variables	Worst in Class ESG	Worst in Class Environmental	Worst in Class Social	Worst in Class Governance
Alpha	0.01204*	0.01218*	0.00878	-0.00144
	(0.00654)	(0.00635)	(0.00649)	(0.00573)
Mkt-RF	1.00997***	0.97976***	0.98617***	1.06985***
	(0.01301)	(0.01330)	(0.01338)	(0.00983)
SMB	0.06176**	0.11506***	0.11397***	0.00012
	(0.02442)	(0.02305)	(0.02278)	(0.01534)
HML	-0.07301***	-0.02821**	-0.00256	-0.08101***
	(0.01266)	(0.01208)	(0.01248)	(0.01037)
MOM	-0.00256	0.00098	-0.00621	-0.03751***
	(0.00899)	(0.00945)	(0.00940)	(0.00849)
R^2	0.9064	0.9094	0.9068	0.9353

Notes to table D2: Robust standard errors are displayed in parentheses. All coefficients and robust standard errors are rounded to five decimals. All R^2 are rounded to four decimals. The number of observations is 4028 for all the portfolios.

* Significant at a 10 percent level, ** Significant at a 5 percent level, *** Significant at a 1 percent level.

Table D3: Changed Market Portfolio (S&P 500) ESG Portfolios 2007-2014 Carhart

Variables	Sector Exclusion	Best in Class ESG	Best in Class Environmental	Best in Class Social	Best in Class Governance
	2007-2014	2007-2014	2007-2014	2007-2014	2007-2014
Alpha	0.00276	-0.00862	-0.00626	-0.01310*	-0.01375
	(0.00406)	(0.00805)	(0.00794)	(0.00796)	(0.00869)
Mkr-RF	0.97602***	0.94188***	0.96141***	0.96057***	1.04314***
	(0.00401)	(0.01284)	(0.01321)	(0.01414)	(0.01425)
SMB	0.04937***	-0.16846***	-0.21249***	-0.22663***	-0.07991***
	(0.00619)	(0.02575)	(0.02473)	(0.02516)	(0.02474)
HML	0.03871***	-0.08038***	0.00655	-0.02312	-0.03160
	(0.00640)	(0.01919)	(0.02127)	(0.02227)	(0.02478)
MOM	-0.02612***	0.00249	-0.00483	-0.01584	-0.11397***
	(0.00477)	(0.01177)	(0.01200)	(0.01208)	(0.01424)
R^2	0.9830	0.9237	0.9318	0.9302	0.9354

Notes to table D3: Robust standard errors are displayed in parentheses. All coefficients and robust standard errors are rounded to five decimals. All R^2 are rounded to four decimals. The number of observations is 2014 for all the portfolios.

* Significant at a 10 percent level, ** Significant at a 5 percent level, *** Significant at a 1 percent level.

Table D4: Changed Market Portfolio (S&P 500) Non-ESG Portfolios 2007-2014 Carhart

Variables	Worst in Class ESG	Worst in Class Environmental	Worst in Class Social	Worst in Class Governance
	2007-2014	2007-2014	2007-2014	2007-2014
Alpha	0.01444	0.01555	0.01176	0.00145
	(0.01100)	(0.01049)	(0.01080)	(0.00948)
Mkr-RF	0.96027***	0.91924***	0.93806***	1.10416***
	(0.02325)	(0.02096)	(0.02231)	(0.01459)
SMB	0.05913	0.11181***	0.09951**	-0.03448
	(0.04295)	(0.03736)	(0.04007)	(0.02513)
HML	-0.00132	0.02037	0.01175	-0.04250
	(0.03505)	(0.03083)	(0.03368)	(0.02615)
MOM	0.01275	-0.01756	-0.00198	-0.05196***
	(0.01785)	(0.01629)	(0.01748)	(0.01534)
R^2	0.8768	0.8840	0.8789	0.9286

Notes to table D4: Robust standard errors are displayed in parentheses. All coefficients and robust standard errors are rounded to five decimals. All R^2 are rounded to four decimals. The number of observations is 2014 for all the portfolios.

* Significant at a 10 percent level, ** Significant at a 5 percent level, *** Significant at a 1 percent level.

Table D5: Changed Market Portfolio (S&P 500) ESG Portfolios 2015-2022 Carhart

Variables	Sector Exclusion	Best in Class ESG	Best in Class Environmental	Best in Class Social	Best in Class Governance
	2015-2022	2015-2022	2015-2022	2015-2022	2015-2022
Alpha	0.00099	-0.00513	-0.00881	-0.00563	0.00724
	(0.00110)	(0.00522)	(0.00608)	(0.00615)	(0.00573)
Mkr-RF	1.00907***	0.96330***	0.91868***	0.92564***	0.98780***
	(0.00160)	(0.00767)	(0.01096)	(0.01255)	(0.00659)
SMB	0.00712**	-0.08665***	-0.13611***	-0.13315***	0.01324
	(0.00289)	(0.01911)	(0.01471)	(0.01509)	(0.01020)
HML	-0.03103***	0.07915***	0.08120***	-0.07779***	0.01140
	(0.00165)	(0.00889)	(0.01132)	(0.01139)	(0.00759)
MOM	0.00200	0.00761	0.00928	-0.00886	0.06790***
	(0.00138)	(0.00813)	(0.01177)	(0.00847)	(0.00566)
R^2	0.9984	0.9590	0.9404	0.9402	0.9544

Notes to table D5: Robust standard errors are displayed in parentheses. All coefficients and robust standard errors are rounded to five decimals. All R^2 are rounded to four decimals. The number of observations is 2014 for all the portfolios.

* Significant at a 10 percent level, ** Significant at a 5 percent level, *** Significant at a 1 percent level.

Table D6: Changed Market Portfolio (S&P 500) Non-ESG Portfolios 2015-2022 Carhart

Variables	**Worst in Class ESG**	**Worst in Class Environmental**	**Worst in Class Social**	**Worst in Class Governance**
	2015-2022	2015-2022	2015-2022	2015-2022
Alpha	0.01022	0.00915	0.00526	-0.00241
	(0.00678)	(0.00677)	(0.00672)	(0.00580)
Mkr-RF	1.05732***	1.03685***	1.04895***	1.00392***
	(0.00970)	(0.01256)	(0.01250)	(141.91)
SMB	0.07239***	0.13646***	0.12640***	0.05475***
	(0.02334)	(0.02451)	(0.02120)	(0.01041)
HML	-0.08136***	-0.02520**	0.02051*	-0.12852***
	(0.01089)	(0.01111)	(0.01086)	(0.00745)
MOM	-0.00627	0.01886*	-0.01059	-0.00821
	(0.00893)	(0.00971)	(0.00860)	(0.00608)
R^2	0.9456	0.9450	0.9454	0.9568

Notes to table D6: Robust standard errors are displayed in parentheses. All coefficients and robust standard errors are rounded to five decimals. All R^2 are rounded to four decimals. The number of observations is 2014 for all the portfolios.

* Significant at a 10 percent level, ** Significant at a 5 percent level, *** Significant at a 1 percent level.