SCHOOL OF ECONOMICS AND MANAGEMENT

The Incorporation of ESG Scores into Factor based Investment Decisions

Does ESG Integration necessarily come with a Financial Trade-off ?

by

Patrick Schumacher

May 2020

Master’s Program in Finance

Supervisor: Hossein Asgharian

Abstract

The present study examines potential financial trade-offs in socially responsible investment strategies. It focuses on a period between September 2009 and November 2019 in Central Europe. While financial performance of an investment is measured by monthly risk-adjusted returns, social responsibility is represented by Sustainalytics ESG scores. First, to analyze the effect of ESG scores on financial stock performance, cross-sectional regressions are applied. To evaluate if the integration of ESG criteria into existing investment strategies significantly affects financial performance, hypotheses tests for equality of intercepts are applied to two self-constructed portfolios that follow the same investment strategy, of which however only one also invests in accordance with ESG criteria. The results of the cross-sectional regression on stock-level signal a negative relation between ESG scores and financial stock performance. A financial trade-off is only observable in one investment strategy while in the remaining strategies sustainable performance improves without significantly reducing financial performance.

Keywords: ESG · Sustainability · Social responsibility · Investment strategies · Financial performance · Integration · Trade-off

Table of Contents

List of Tables

List of Figures

1 Introduction

In recent years, environmental, social and governance (ESG) investments have become increasingly popular for investors around the world. While the climate change became reality, the value of an investment is no longer measured solely in terms of risk and return, but also in terms of its positive impact on society. In this context, socially responsible investing (SRI) in particular has caused a lot attention among investors. According to a 2018 survey of the U.S. Forum for Sustainable and Responsible Investment, more than 25% of all assets under management in the U.S. are invested socially responsible (US SIF, 2019). The purpose of this paper is to conduct an empirical study that investigates the financial impact of social responsibility on existing investment processes and decisions in Central Europe. By using a variety of different approaches, it aims to achieve reliable results that further contribute to the research fields of sustainable finances and ESG.

In general, SRI actively selects investments according to prespecified guidelines which for example can be driven by personal values, climate goals or other factors. In order to screen a whole investment universe for potential investments that meet those guidelines, ESG scores are taken into account. ESG looks into a firm's environmental practices, its social relations and its corporate governance and finally rates the business according to them. The higher a score of a respective firm compared to other peers from the same industry, the more likely that this firm will meet the given guidelines. Overall, it can be stated that socially responsible investors besides financial performance are also concerned about a non-financial dimension of performance which is represented by ESG scores (Galema, Plantinga & Scholtens, 2008). Renneboog, Ter Horst and Zhang (2008a) even go one step further by stating that if investors derive non-financial utility from socially responsible investments, they consequently care less about financial performance than traditional investors would do. This raises the question whether a trade-off between financial and non-financial dimension of performance in socially responsible investments exists. In a 2017 survey carried out by the Morgan Stanley Institute for Sustainable Investing, 53% of all investors asked, believe that SRI comes with a financial trade-off (Morgan Stanley, 2017). Over recent years, an extensive amount of empirical research has been carried out in order to evaluate the performance of sustainable investments in comparison to traditional ones. While studies like Hamilton, Joe and Statman (1993), Goldreyer and Diltz (1999), Statman (2000) and Bello (2005) show that financial performance of socially responsible and traditional investments in the U.S. are not significantly different from each other, studies like Hong and Kacperczyk (2007) or Renneboog, Ter Horst and Zhang (2008b) report an underperformance of SRI funds in Europe and Asia. On the other side, studies like Kempf and Osthoff (2007) report significant positive risk-adjusted returns for U.S. long-short SRI portfolios. Overall, a recent meta-analysis of over 2000 empirical studies by Friede, Busch and Bassen (2015) summarizes that 15.5% of all studies show positive risk-adjusted returns while 11% show negative ones. The remaining 73.5% show no change in financial performance or report mixed results. Based on the multitude of different findings from preceding literature, this paper picks up the relationship between ESG scores and stock returns and investigates whether socially responsible investing can improve financial performance of an investor's portfolio without failing to meet non-financial SRI guidelines. However, in contrast to most previous studies where portfolios are constructed solely on the basis of ESG scores, a more feasible approach is chosen. Since in reality, fund managers rather use ESG scores as an add-on to their existing investment process (van Duuren, Plantinga & Scholtens, 2016), this paper aims to investigate a potential financial trade-off in socially responsible investments by incorporating ESG scores into existing investment strategies. By doing this on a recent data set, results will most likely be clearer than in past studies since awareness of ESG measures constantly increases among investors.

The research approach of this paper comes in two parts. While the first part examines the general link between ESG scores and financial stock performance on a stock-by-stock level, the second part actually investigates financial performance on a portfolio-level. Generally, this paper uses Sustainalytics ESG scores obtained from Bloomberg as a proxy for social responsibility. The main advantage of using Sustainalytics as a rating provider is that ESG scores are reported monthly both as an overall score as well as environmental, social and governance pillar scores. Consistent with Galema, Plantinga and Scholtens (2008) and Mănescu (2011), the first part of the research approach mainly covers two methods. First, a cross-sectional regression to investigate the general relation between ESG scores and stock returns and second a pooled regression to explain potential industry and country specific effects. Subsequently, the second part deals with the performance of self-constructed SRI portfolios. In their research paper, Kempf and Osthoff (2007) construct long-short SRI portfolios based on the portfolio approach established by Fama and French (1993) that invest in the 10% best SRI stocks and short sell the 10% worst SRI stocks. This approach uses ESG scores as the basis of an investment strategy, resulting in so called pure ESG portfolios. While non-financial performance is ensured by investing in stocks with high ESG scores and short selling stocks with low ESG scores, the aim is to measure financial performance represented by risk-adjusted returns. Comparing this approach with the construction of portfolios in this paper, ESG scores are rather treated as an add-on to existing investment processes. Using a similar framework, Melas, Nagy and Kulkarni (2016) analyze the impact of ESG integration into different factor investment strategies. Based on their results, ESG integration tends to improve both risk-adjusted portfolio performance as well as the ESG level of a strategy without significantly affecting the target factor exposure. For this paper, the integration effect of ESG scores into existing investment strategies is measured by comparing risk-adjusted returns of two self-constructed portfolios that follow the same investment strategy, of which only one however also invests in accordance with ESG criteria. While pre-integration portfolios are constructed according to value, size, momentum, low volatility, quality and growth investment strategies, post-integration portfolios are constructed by additional consideration of ESG criteria in each strategy. Finally, by testing if risk-adjusted returns significantly differ between two portfolios representing the same strategy, a potential financial trade-off in socially responsible investments can be investigated. Additionally, for the sake of completeness, pure ESG portfolios are constructed and analyzed in line with the approaches from previous literature.

Overall, results from the cross-sectional regression show a significant negative relation between ESG scores and financial stock performance which indicates an outperformance of firms with a low degree of social responsibility. However, time-series regressions on portfolio-level show that both pure ESG portfolios as well as those portfolios where ESG serves as an add-on deliver significant positive risk-adjusted returns. Furthermore, in the majority of strategies, the integration of ESG scores into existing processes does not significantly worsens financial stock performance. Therefore, the conclusion can be drawn that the incorporation of ESG scores into existing investment strategies does not come with a financial trade-off, even though cross-sectional regressions indicate a negative relationship.

The remainder of this paper is structured as follows: Chapter 2 gives an overview of the existing literature related to financial trade-offs in socially responsible investments. Furthermore, it summarizes the main findings related to ESG effects on financial performance of firms. Chapter 3 develops this paper's hypotheses, chapter 4 discusses the underlying data set and chapter 5 explains the methodological approaches used. Chapter 6 presents and discusses the results of all methods applied, whereas chapter 7 finally summarizes and concludes.

2 Literature Review

This chapter serves to give a general overview of the existing literature related to financial trade-offs in socially responsible investments. As already outlined in the introduction to this paper, several studies have dealt with the relationship between socially responsible investments and financial stock performance and mixed evidence was found. While the majority found positive correlations between ESG and financial performance, a not inconsiderable amount of studies also found negative or nonexistent correlations (Giese, Lee, Melas, Nagy & Nishikawa, 2019). In the following, this relationship is first examined separately on the basis of theoretical and empirical models. Subsequently, basic concepts and assumptions of both types of models are compared with each other in order to eliminate potential shortfalls of either one of them. Finally, a few selected articles which deliver significant results are analyzed in greater detail to supplement the justification for the methodological approaches that will be used later in this paper. Consistent with the two step approach introduced in chapter one, the first couple of papers deal with the general link between ESG scores and financial performance of firms while the remaining papers deal with ESG integration into investment strategies.

2.1 Theoretical Approach

In Renneboog, Ter Horst and Zhang (2008a), a first critical review of some literature on socially responsible investing is given. In order to understand the relationship between SRI performance and financial performance of a company, one first has to understand the background of why companies may act socially responsible. While Finance textbooks argue that companies always should only care about maximizing the value of shareholders' equity (Jensen, 2004), SRI investors contradict this statement. They also promote socially and environmentally sound corporate behavior, which represents a maximization of the social value (Renneboog, Ter Horst & Zhang, 2008a). If now shareholder value maximization of a company does not equal the maximization of social value for all stakeholders, the company faces a trade-off decision. In this situation, each company has to decide whether or not it is worth to act socially responsible. Heal (2005) states that socially responsible behavior minimizes potential conflicts between a company and society, which comes in line with a reduction of costs of conflicts and thus maximizes shareholder value. Another argument in favor of social responsibility is given by Allen, Carletti and Marquez (2007) who state that stakeholder-oriented firms have higher firm values compared to shareholder-oriented firms because they face a reduced competition in selling ethical brands. Overall, Renneboog, Ter Horst and Zhang (2008a) summarize that rational firms may voluntarily act socially responsible due to less competition and an improvement in reputation. These factors can increase firm value and that is what every company is looking for.

Coming back to the relationship between SRI and financial performance, theoretical studies cite differences in demand for different types of stocks as a main reason for variations in financial performance (Galema, Plantinga & Scholtens, 2008). While the differences in demand can result from incomplete information about a firm's expected return, its variance and covariance with other stocks (Merton, 1987), from investor preferences for socially responsible companies (Heinkel, Krause & Zechner, 2001) or other non-financial performance characteristics (Fama and French, 2007), they all result in discrepancies in underlying stock prices. One explanation is that while socially responsible stocks face excess demand what makes them overpriced, socially irresponsible stocks face a shortage of demand resulting in underpricing (Galema, Plantinga & Scholtens, 2008). If an investor now decides to buy this underpriced stock, his risk sharing opportunities are limited which in turn requires him to ask for a return premium (Merton, 1987). Since SRI investors do not face risk sharing problems, theoretical studies like Dam (2008) conclude that socially responsible companies generate lower returns then their irresponsible competitors.

2.2 Empirical Approach

While theoretical explanations argue for financial trade-offs in socially responsible investments, empirical literature like Hong and Kacperczyk (2007), Kempf and Osthoff (2007) or Statman and Glushkov (2009) comes to another result. All these studies find that a firm's social responsibility either results in positive abnormal returns or in no significant changes of financial performance. Similar to the theoretical approach, Derwall, Guenster, Bauer and Koedijk (2005) argue that mispricing of socially responsible stocks is the cause for these positive abnormal returns. Investors seem to underestimate social responsibility or overestimate its cost. However, while the theoretical literature uses mispricing as an explanation for negative abnormal returns, empirical literature like Derwall et al. (2005) argues in the opposite direction. According to Galema, Plantinga and Scholtens (2008), the contradiction between theoretical and empirical literature results from a misinterpretation of risk-adjusted financial performance in most empirical studies. In contrast to theoretical models, empirical models implement risk adjustments by controlling for several risk factors in a multi-factor model like those suggested by Fama and French (1992). In these multi-factor models, an intercept that is not significantly different from zero is proof enough that a respective asset is priced according to a respective asset pricing model. These models however ignore a characteristic that is considered in theoretical models, namely the excess demand for SRI stocks. Overpriced SRI stocks typically have a lower book-to-market (BTM) ratio than their underpriced, socially irresponsible competitors. If a SRI stock and a non-SRI stock have the same level of risk but different book-to-market ratios, Fama and French multi-factor models imply that BTM factors do not affect the risk profile of both firms. This drawback however seems to be wrong since the trade-off between SRI performance and financial performance is at least partly driven by the BTM ratio (Galema, Plantinga & Scholtens, 2008). For this reason, the authors recommend to investigate both, the BTM ratio as well as a Fama and French multi-factor model in order to draw consistent conclusions for financial performance of socially responsible investments. A second problem that may explains the contradiction between theoretical and empirical literature is the aggregation over different pillars towards one overall ESG score that expresses social responsibility. A research paper by Scholtens and Zhou (2008) for example states that environmental improvements of a firm’s business increase its expected return while social improvements have the opposite effect. Therefore, the investigation of different dimensions of ESG scores is necessary to get empirical results that are in line with theoretical models. In a next step, this paper takes a closer look at several empirical studies that pick up and overcome the potential problems mentioned above. While the first couple of papers deal with the general link between ESG scores and financial stock performance of firms, the remaining papers deal with ESG integration into factor investment strategies.

2.3 ESG Scores and Financial Performance

During recent years, several different approaches were used to analyze the relationship between social responsibility and financial performance of firms. Before going into more detail about these individual approaches, it is worth mentioning that a main difference between all research in the area of ESG is the measurement of sustainability. While some literature like Cohen, Fenn and Konar (1997) or Konar and Cohen (2001) only focus on single sustainability criteria like measures of pollution or other individual ESG pillars, more recent literature like Galema, Plantinga and Scholtens (2008) or Mănescu (2011) consider an overall sustainability screening represented by aggregated ESG scores or other corporate social responsibility (CSR) ratings as their benchmark to quantify SRI. In order to prevent the previously explained potential contradiction between theoretical and empirical models, this paper uses a data set of aggregated ESG scores that will first be investigated from its three individual pillar dimensions before using the overall score to draw a conclusion.

Consistent with the order of methodological approaches that will be applied in this paper, previous literature using either a cross-sectional regression approach or a portfolio analysis approach is reviewed first. While results of a cross-sectional analysis can make a statement about the general relationship between ESG scores and financial stock performance on a stock-level, portfolio analysis goes a little further into detail and examines the effect of ESG scores on risk-adjusted portfolio returns using time-series regressions. Finally, although it is not used in this paper, the event study approach is shortly reviewed for the sake of completeness.

2.3.1 Cross-sectional Regression Approach

The cross-sectional regression approach goes back to the time of Fama and French (1992). In their paper, the authors state that systematic market risk, which is expressed as beta, is not sufficient to explain cross-sectional variations in stock returns. Instead of using only beta, they found that cross-sectional variations are mainly captured by the size and the book-to-market ratio of a company. Based on this framework, every potential economic or corporate key figure can be implemented as an explanatory control variable into the model in order to investigate if it helps to explain cross-sectional variations in stock returns. In their follow-up paper, Fama and French (1993) therefore expand Sharpe and Lintner's Capital Asset Pricing Model (CAPM) by two additional risk factors that represent both size and book-to-market ratios. The resulting so-called Fama and French three-factor model still serves as the basis model for many researchers in order to investigate the effect of ESG scores on stock returns. While for example Ziegler, Schröder and Rennings (2007) expand Fama and French's original three-factor model only by a sustainability measure, Mănescu (2011) further takes the Carhart (1997) momentum factor as an additional risk factor into account. In contrast to these models, Galema, Plantinga and Scholtens (2008) do not make use of Fama and French's risk factors but use the logarithm of corporate key figures in their cross-sectional regression.

In Galema, Plantinga and Scholtens (2008), the authors apply the cross-sectional regression approach to investigate the relationship between six different SRI variables and excess returns of individual stocks. For this purpose, they use a two-step Fama-MacBeth (1973) regression where market risk betas are first estimated in a time-series regression of individual stocks on market returns. The second step then consists of the cross-section regression itself. Here, the time-series beta estimates are regressed as explanatory variables together with additional control variables and the SRI variable against excess stock returns in order to analyze the investigated relationship. For their model, the authors use the following control variables: the natural logarithm of each firm's market capitalization lagged by one month, the logarithm of the book-to-market ratio lagged by one month as well as the firm's average return over the last 12 month and the monthly average of daily turnover lagged by one month. Additionally, the SRI variable is also lagged by one month since market participants do not shift their investment decisions immediately after updated SRI variables are announced. While SRI variables are obtained from KLD Research & Analytics, the remaining data on financial performance measures are taken from Thomson Reuters Datastream. After running their regression, Galema, Plantinga and Scholtens (2008) found that, consistent with Fama and French's original findings from 1992, betas can only insignificantly describe expected returns in their framework. Furthermore, turnover also delivers insignificant results. On the other hand, book-to-market ratios and market capitalizations both have a negative effect on excess returns. However, while BTM ratios are significant at a 1% level, the size of a firm is only significant at a 10% level. Turning the focus on the actual relationship of interest, the authors found that only employee relations, which can be categorized in the social ESG pillar, have significant positive effects on a company's excess return. The remaining SRI variables do not show any significance in explaining the cross-section of excess stock returns.

Galema, Plantinga and Scholtens (2008) further perform a book-to-market cross-sectional regression whose results are in line with the theoretical model of demand differences between socially responsible and irresponsible stocks. In the regression, the authors investigate the impact of ESG scores on the value of a firm. By using the logarithm of book-to-market ratios as the dependent variable and ESG scores together with several control variables that are well known as BTM predictors as explanatory variables, the regression states that ESG scores have a significant impact in the BTM ratio. While the environmental and social pillars have a negative impact on the BTM ratio, the governance pillar's impact is positive. Thus, the authors conclude that ESG scores affect stock returns by changing the book-to-market ratio of a firm which leads to an adaption of valuation and an adjustment in returns.

In contrast to the approach in Galema, Plantinga and Scholtens (2008), Mănescu (2011) conducts a cross-sectional regression while controlling for Fama and French risk factors instead of logarithms of corporate key figures. However, both papers use the same ESG scores obtained from KLD Research & Analytics for the same period of time between 1992 and 2006, what makes results of both methods comparable to each other. In conjunction of the Fama-MacBeth regression, Mănescu (2011) also applies a time-series regression first in order to obtain market risk betas. The subsequent month-by-month cross-sectional regression is then conducted with beta as well as with size, value and momentum factors as control variables. The central economic interpretation behind this framework is that the expected return of a stock should be high if it has large exposure to factors that carry risk premium (Mănescu, 2011). Again, all risk factors as well as the ESG variables are lagged by one month. While market capitalizations represent the size risk factor, book-to-market ratios represent the value risk factor. Additionally, the momentum risk factor is represented by average returns between the month t – 2 and t – 12. Overall, Mănescu (2011) states similar results compared to Galema, Plantinga and Scholtens (2008). While both the BTM and the momentum factor significantly explain the cross-section of returns, beta and the size factor do not. In terms of ESG variables, only the community relations factor, which again can be categorized in the social ESG pillar, shows significant results. Looking at aggregated ESG scores, no statistically significant effect on the cross-section of stock returns is observable. According to Mănescu (2011), this result is consistent with previous studies like Derwall and Verwijmeren (2007) and needs to be considered in further research on this topic.

Besides the general relationship between ESG scores and financial stock performance of firms, Mănescu (2011) further analyses industry specific effects of ESG scores. According to the author, it is necessary to control for industry specific effects in order to avoid potential misinterpretations of the investigated relationship. Previous research like Belu (2009) supports this statement by showing that under consideration of aggregated ESG scores, ESG performance differs significantly between different industries. In order to account for potential confounding effects, Mănescu (2011) further adds industry dummies to its four-factor model.

2.3.2 Portfolio Analysis Approach

The portfolio analysis approach is most commonly based on factor mimicking portfolios and closely related to the construction of size and value factors which Fama and French (1993) use for their three-factor model. Mimicking portfolios have a unit exposure to a background factor that they represent and the expected return of the portfolio equals the risk premium of the factor (Asgharian, 2004). A common approach to construct such factor mimicking portfolios is to take a long position in assets with high factor loadings and a short position in assets with low factor loadings. Previous literature follows this approach in different ways. While some studies construct ESG portfolios using single sustainability measures like pollution (Cohen, Fenn & Konar, 1997) as their factor to decide which assets they buy and which they short-sell, other more recent literature like Kempf and Osthoff (2007) and Mănescu (2011) construct factor mimicking portfolios based on ESG scores or other aggregated sustainability ratings. After the respective portfolio is constructed, its performance can be evaluated using monthly portfolio returns as dependent variable in an OLS time-series regression to test for risk-adjusted returns.

In their research paper, Kempf and Osthoff (2007) analyze the question whether investors can increase financial stock performance by incorporating social responsibility screens into their investment process. Therefore, they implement different trading strategies of buying stocks with high ratings and short selling those with low ratings. Furthermore, they construct two additional portfolios where the first one represents stocks with high SRI ratings and the second one stocks with low SRI ratings. The authors find that over a period from 1992 until 2004, the portfolio representing stocks with high SRI ratings performs better than its comparison portfolio with low ratings. After regressing excess returns of their long-short SRI portfolio in a four-factor model using the original Fama and French three-factor model plus Carhart's momentum factor, Kempf and Osthoff (2007) further investigate a positive abnormal return of 8.7% per year. This finding proves their hypothesis of an increased financial performance after incorporating social responsibility screens into the investment process. A different approach of how to use factor mimicking portfolios is applied in Mănescu (2011). Here, the author applies a method developed by Charoenrook and Conrad (2005) in order to test if ESG scores can be treated as a risk factor. The test is based on the relationship between the conditional mean and the conditional return variance of a factor mimicking portfolio, which have to be linearly related to each other for ESG scores to be a risk factor. Mănescu (2011) uses a GARCH-in-mean model and constructs its factor mimicking portfolio by going long in the 30% of stocks with the highest rating and selling the 30% of stocks with the lowest rating. The resulting Low-Sustainability minus High-Sustainability (LMH) risk factor can then be used in a time-series regression to investigate whether ESG consideration as the basis of an investment process can improve financial performance by delivering positive risk-adjusted returns.

2.3.3 Event Study Approach

Even though this approach is not applied in this paper, previous event studies deliver enlightening insights about the reaction of firms' financial performance on sustainability related events, which is why this section shortly reviews some of the most popular event studies.

Two of the first studies examining announcement effects of sustainability related events were Hamilton (1995) and Konar and Cohen (1997). In both event study, the authors investigate the effect of environmental pollution information on the stock price of affected firms. The outcome of these studies both state significant negative abnormal returns after the announcement of environmental pollution. In line with these findings but analyzed from an opposite point of view, Klassen and McLaughlin (1996) perform an event study that concludes positive abnormal stock returns resulting from positive environmental announcements. Even though these studies only consider the environmental pillar of today's available aggregated ESG scores, findings are in line with those resulting from other methods like the portfolio analysis approach. However, results have to be treated with caution. McWilliams, Siegel and Teoh (1999) criticize that event studies only investigate short-run impacts and that already small changes in the research design can lead to insignificant results. This is why an event study approach is not applied in this paper.

2.4 ESG Integration into Factor Investment Strategies

Up to this point in the review, existing literature only focuses on financial trade-offs in socially responsible investments. For this reason, long-short ESG portfolios were constructed in order to investigate potential risk-adjusted abnormal returns. However, as stated by van Duuren, Plantinga and Scholtens (2016), this way of applying ESG scores as the basis of an investment decision generally does not correspond to reality. According to the authors, ESG scores rather act as an add-on to existing investment strategies than as an independent factor for long-short ESG strategies. Therefore, this part of the literature review focuses on the integration of ESG scores into existing investment strategies or so called investment styles.

Considering the integration of ESG scores into value, growth and momentum strategies, Kaiser (2020) is one of the latest research papers which demonstrates that investors can increase sustainability performance of their portfolio without sacrificing any financial performance. In his paper, the author first applies multi-dimensional passive screens based on the Morningstar style box in order to identify and rank stocks consistent with their appropriate style. The

Morningstar style box replaces common style signals like for example the book-to-market ratio for value stocks with a group of five other value-rating components that are both forward and backward looking. Furthermore, he categorizes all stocks according to their aggregated Asset4 ESG scores obtained from Thomson Reuters between 2002 and 2015. In order to combine both the style and ESG ranking of a stock into one combined measure, Kaiser (2020) calculates a combined average rank across both measures, which is then used to sort all stocks into size-weighted quintile portfolios. Instead of constructing long-short zero-investment portfolios as recommended by Asness, Moskowitz and Pedersen (2013), Kaiser (2020) constructs long-only portfolios "for the benefit of a broader application across investment managers" (p.12). The resulting long-only portfolios that accommodate the combined average rank are then rebalanced every year and performance is compared to one-dimensional style portfolios as well as to a long-only ESG portfolio. This procedure is performed twice, once for European and once for American stocks. Looking at the key results of this research paper, Kaiser (2020) finds evidence for an improvement of risk-adjusted returns in all three strategies after the integration of ESG scores. While the growth strategy delivers an improvement of risk-adjusted returns both for European and American stocks, value and momentum strategies both increase their risk-adjusted returns at least considering American stocks. Furthermore, Kaiser (2020) states that both European and American investors can increase the sustainability rating of their portfolios by at least 25% without a statistically significant decrease in financial performance.

The question that remains unanswered is the explicit impact of ESG integration from a strategic point of view. In their research paper, Melas, Nagy and Kulkarni (2016) address this issue by examining exactly this impact for six different factor strategies, namely low volatility, quality, yield, value, size and momentum. In a first step, the authors analyze the relationship between ESG scores and all respective factors by measuring cross-sectional correlations between factor exposures and ESG scores on a stock-by-stock level. According to Melas, Nagy and Kulkarni (2016), the resulting low level of average correlations make ESG scores to an independent source of new information. However, some intuitive and statistically significant relationships can still be found, especially in positive ESG correlations with size, quality and low volatility factors. After having a general idea about the correlation between factors and ESG scores, Melas, Nagy and Kulkarni (2016) assess the impact of ESG integration into all factor strategies by applying a portfolio construction framework. Starting with the results of ESG integration into a low volatility factor strategy, the authors find only an insignificant increase in realized volatility. They state that it is therefore possible "to improve the ESG ratings of minimum volatility strategies without a significant impact on the risk reduction properties and overall characteristics of the strategy" (Melas, Nagy & Kulkarni, 2016, p.13). Moving on to quality factor strategies, the authors report that stocks with a high exposure to quality factors also tend to have above-average ESG scores which leads superior financial performance for portfolios that combine quality and ESG measures. Turning the focus to those strategies where factors had zero or negative correlation with ESG scores, the results tend to be mixed. While value strategies tend to increase their information ratio and thus their risk-adjusted return after the integration of ESG measures, momentum strategies experience a minor decrease in risk-adjusted performance. Overall, Melas, Nagy and Kulkarni (2016) find evidence that both defensively oriented as well as dynamic factor strategies can increase their sustainability performance without or with only modest adjustments in their primary investment objectives.

3 Hypotheses Development

After reviewing previous literature in the field of sustainable finance and ESG scores, this chapter serves the actual hypotheses development. More precisely, the main hypothesis of no financial trade-off in socially responsible investments will be divided into three sub-hypotheses. As already mentioned before, focus is put on the investigation whether socially responsible investing can improve financial performance of an investor's portfolio without failing to meet non-financial SRI guidelines. Starting with the cross-sectional relation between ESG and financial stock performance, empirical evidence suggests that stocks with high ESG scores earn positive abnormal returns (Derwall et al., 2005; Statman and Glushkov, 2009). If this connection turns to be true, it supports the main hypothesis of no financial trade-off in socially responsible investments since these investments use ESG scores as a proxy for social responsibility. Therefore, the first sub-hypothesis is as follows:

Hypothesis 1: Firms with high ESG scores outperform their lower ranked peers

Turning the focus away from stock-level relationship towards a portfolio-level, even a negative relation between ESG scores and financial stock performance does not necessarily need to result in a financial trade-off. This is if the consideration of social responsibility in an investment process does not significantly worsen returns but rather increases non-financial performance by investing socially responsible. Therefore, the second sub-hypothesis targets the financial performance of SRI portfolios. By comparing risk-adjusted returns of portfolios with and without ESG consideration, it is formulated as follows:

Hypothesis 2: The integration of ESG scores into existing investment strategies does not reduce financial performance but increases sustainable performance

In light of this hypothesis and the fact that most preceding literature uses ESG as the basis of an investment decision rather than as an add-on, a third supplementary sub-hypothesis is added:

Hypothesis 3: Pure ESG portfolios deliver statistically significant positive risk-adjusted returns

Overall, all sub-hypotheses serve the aim to shed light on the main hypothesis of no financial trade-off in socially responsible investments and will be answered in this context.

4 Data and Variables

This chapter briefly discusses the underlying data and variables used to test the main hypothesis of no financial trade-off in socially responsible investments. As already outlined in chapter 1, focus is put on financial sustainability in Central Europe, which is why the STOXX Europe 600 index (ISIN: EU0009658202) serves as the underlying investment universe for this analysis. With a fixed number of 600 companies included, the STOXX Europe 600 index represents small, mid and large capitalization firms across 17 countries of the European region. Its broad coverage enables both the investigation of industry and country specific effects and is thus the optimal choice for this paper. In order to measure both sustainable performance as well as financial performance for stocks included in the index, two main types of data are necessary. First, financial data is gathered from Thomson Reuters Datastream. This data mainly consists of stock prices and further fundamentals that are used as control variables in cross-sectional regressions. Second, sustainability data is obtained from Sustainalytics, an ESG database available in Bloomberg. Sustainalytics is an ESG rating provider that covers more than 11,000 companies across 138 sub-industries. Its main advantage compared to other ESG rating providers is that ESG scores are reported monthly instead of only yearly both as an aggregated score as well as environmental, social and governance pillar scores. The reporting of a monthly score has the advantage of a more detailed time-series component, as adjustments in ESG scores can also be made during the year. In general, Sustainalytics ESG scores in Bloomberg are available retroactive until September 2009. However, not all stocks included in the STOXX Europe 600 index have frequently reported ESG scores back till then for different reasons. While some firms did not report sustainability measures because of a lack of necessity in the past, others simply did not exist yet. Therefore, the cross-sectional scope of this paper has to be adjusted downwards. For a firm to be included in the analyzed sample, both financial data as well as ESG scores have to be available for at least 12 months in a row. This results in an overall sample of 289 firms with a maximum of 123 monthly observations per firm between September 2009 and November 2019. Out of these 289 firms, 84 represent a small capitalization, 96 a mid capitalization and 109 a large capitalization. The sample covers 40 different niche industries from 12 different countries and will be used for the cross-sectional regression of stock returns. For the construction of pure ESG and other strategy portfolios in step two of the research approach, this sample then has to be further modified according to two criteria. First, monthly ESG scores have to be available without exceptions from January 2014 onwards. The year 2014 is chosen as a trade-off between keeping both the number of firms as well as the time-series dimension relatively high. Second, stock prices have to be available from December 2012 onwards. This is due to the fact that for the construction of momentum factors, stock returns are already needed for the preceding 12 months, which makes stock prices for December 2012 the first ones needed. Instead of covering all 289 firms, the modified sample consists of 44 small caps, 54 mid caps and 95 large caps, which results in a total of 193 stocks analyzed between January 2014 and November 2019. In the following sections, the individual data from both samples are described in more detail and their respective use is explained.

4.1 Financial Data

Looking at financial performance, the most important measure for the following analysis is the risk-adjusted return of a company. Therefore, monthly stock returns for each asset i, are calculated as the percentage difference between logarithmic prices at the beginning of two consecutive months:

$$r_{i,t} = \ln\left(P_{i,t}\right) - \ln\left(P_{i,t-1}\right) \tag{1}$$

where $P_{i,t}$ indicates the price of stock $i = 1, \dots, N$ at time $t = 1, \dots, T$. Subsequently, in order to obtain excess returns, the three month EURIBOR serves as the risk free interest rate that will be subtracted from monthly stock returns $r_{i,t}$. Furthermore, risk included in asset i is expressed as its volatility and calculated as the standard deviation of monthly stock returns.

Additionally, for the measurement of risk-adjusted returns using a Fama and French four-factor model, monthly size, value and momentum risk factors are collected for the European market from the Kenneth R. French Data Library. Their individual calculation will further be explained in chapter 5 of this paper when it comes to the construction of factor mimicking portfolios. Even though knowing these risk factors are not constructed based on the same sample of firms that is used for this paper's analysis, they are considered appropriate due to their broad coverage of the overall European market.

For the cross-sectional regression, calculated stock returns from above are regressed against lagged ESG scores together with additional control variables that represent specific firm characteristics. For this paper, control variables are the market capitalization of a firm, its book-to-market ratio as well as its leverage ratio. While market capitalization can be obtained directly from Thomson Reuters Datastream, the remaining two control variables are calculated as follows. First, for the leverage ratio of a firm, its total reported debt is divided by the market value of equity. Second, for the book-to-market ratio, the book value per share is divided by the share price that day.

Last but not least, some more financial data is needed for the construction of portfolios that represent existing investment strategies in the second part of this paper's research approach. While size, value, momentum and low volatility portfolios will be constructed based on variables already explained above, two additional measures of style exposure are needed for growth and quality strategies. For the growth stocks investment style, earnings per share (EPS) growth rates are used as sorting variables. Therefore, yearly EPS are obtained for every firm and average growth rates over three years are calculated. For the quality stocks investment style, sorting variable is the return on equity (RoE) of a firm. Here, the variable does not need to be further adjusted. The general procedure of how all investment style portfolios are constructed will be further explained in chapter 5 of this paper. Now that all necessary financial data has been explained, the next section looks at ESG scores as a measure for sustainability.

4.2 Sustainability Data

As already mentioned in the introduction to this paper, the main research question deals with a potential financial trade-off in socially responsible investments. While the last section mainly focused on the financial component of this question, this section deals with a measure that expresses the degree social responsibility within corporations. The talk is of ESG scores. Due to the increasing demand for socially responsible investments in recent years, a multitude of rating agencies started to make a business with the reporting of ESG scores. However, rating procedures and frequencies often vary over different providers, making their ESG reports difficult to compare. Tailored to the requirements of this analysis, Sustainalytics ESG scores obtained from Bloomberg are selected as the reference measure for social responsibility and sustainable performance. Sustainalytics is a global leader in ESG ratings with over 25 years of experience. With more than 220 indicators, Sustainalytics ESG data on environmental, social and corporate governance topics offer a broad coverage of information based on a high-quality research approach. While the environmental pillar mainly consists of information on emissions and waste production, the social pillar is composed of human rights and employee related topics. Finally, the corporate governance pillar covers information on the company's management and its shareholders. Both the aggregated as well as the three ESG pillar scores take on values between 0 and 100 and can therefore be easily compared. While a score close to 100 attests to very sustainable corporate behavior, a value close to 0 means the opposite. Consistent with Kaiser (2020), this paper explicitly includes all three ESG pillar scores in its research approach rather than defining sustainable performance solely on the basis of the environmental and social pillar. Whilst retail investors primarily focus on environmental and social pillars (Berry and Junkus, 2013), professional investors also take the governance pillar into account (van Duuren, Plantinga & Scholtens, 2016). Since this paper's aim is to draw a conclusion as general as possible, all three pillars are included. Therefore, the equally weighted aggregated ESG score is used for the construction of all ESG portfolios later in chapter 5.

Looking at the overall sample of 289 firms, Sustainalytics assigns a peer industry to every firm included. However, since some of these industries are very small niche industries with only a few firms included, a broader classification is further used in this paper. Therefore, some niche industries are grouped together into one superordinate group. Overall, this shrinks the original 40 niche industries to 13 broader industries which are then used as dummy variables in a pooled regression. The final allocation of aggregated ESG scores to industries and countries can be found in *Table 1* below.

Table 1: Allocation of aggregated ESG scores to industries and countries

Industry	*Number of obs.*	*ESG average*
Utilities	36	73.7
Financial	49	65.3
Real Estate	19	60.5
Transportation	23	65.7
Automobile	14	70.7
IT	48	63.7
Construction	12	65.5
Food	16	66.9
Machinery	11	64.3
Pharma & Chemical	28	64.4
Healthcare	10	62.3
Textiles & Apparel	10	65.5
Others	23	67.5
Overall	*289*	

Country	*Number of obs.*	*ESG average*
Austria	8	62.7
Belgium	16	60.5
Finland	16	70.7
France	82	66.3
Germany	70	64.0
Ireland	8	58.9
Italy	27	68.9
Luxembourg	7	63.4
Netherlands	29	68.5
Spain	25	69.7
Switzerland	1	84.1
United Kingdom	7	72.9
Overall	*289*	

At this point it is worth mentioning that some industry and country specific effects resulting from the pooled regression performed later in this paper might be insignificant due to the limited number of firms representing an industry or country. However, core industries like Utilities, Financials and Information Technologies (IT) are represented by a sufficient number of firms and should therefore deliver trustworthy results. Before looking at these results in chapter 6, chapter 5 next explains the methodological approaches that will be carried out through the rest of this paper.

5 Methodology

In the following chapter, the two methodological approaches used in this paper are explained in further detail. First, cross-sectional as well as pooled regressions on a stock-level are applied to understand the general relation between ESG scores and financial performance. Answering the first sub-hypothesis, this procedure aims to investigate if firms with higher ESG scores also generate higher stock returns compared to firms with lower ESG scores. Assuming this positive relationship holds, ESG scores could be used as a signal for future investment decisions. However, since a cross-sectional regression is limited to a stock-by-stock analysis, this paper's second approach involves a time-series analysis of stock portfolios that are constructed according to ESG criteria. Using a Fama and French (1993) three-factor model extended by Carhart's (1997) momentum factor, risk-adjusted returns for these portfolios can be measured and conclusions regarding financial performance of socially responsible investment strategies can be drawn. Knowing that in practice ESG criteria are generally not the basis of an investment strategy but rather act as an add-on to it (van Duuren, Plantinga & Scholtens, 2016), this paper investigates both types of ESG integration into an investment process. First, stock portfolios are constructed solely on the basis of ESG scores, while afterwards investment strategies are replicated both with and without ESG integration. In order to test for significant differences in financial performance of two portfolios that represent the same investment strategy, once with and once without ESG integration, hypothesis testing for equality of intercepts will be applied.

5.1 Cross-sectional Regression on Stock-level

In a first step, cross-sectional regressions on a stock-level are applied to understand the general relationship between ESG scores and financial stock performance. This procedure is based on Fama and French's research paper from 1992, dealing with the cross-section of expected stock returns. While up to this point in time Sharpe's (1964) Capital Asset Pricing Model (CAPM) stated that cross-sectional differences in stock returns are solely explained by market betas, Fama and French found evidence that contradict this statement. In their research paper, the authors only investigated a weak explanatory power of market betas but therefore found high sensitivities in stock returns to a firm's market value, its book-to-market ratio as well as its leverage and earnings-to-price ratios. Hence, Fama and French (1992) implemented a cross-sectional regression model that in addition to market betas also include those firm specific variables listed above. However, more importantly, they show that any firm specific variable can be integrated into this model to analyze its influence on stock returns. Looking for example at Galema, Plantinga and Scholtens (2008), the authors use a regression model which, next to a variety of other control variables, contains firm specific SRI scores as an independent variable to obtain the effect of social responsibility on stock performance.

In this paper, a similar approach is followed. While monthly stock returns as calculated in equation 1 are used as dependent variable, monthly Sustainalytics ESG scores serve as the independent variable of interest to capture the effect of social responsibility on financial stock performance. Furthermore, following Fama and French (1992), a firm's market value, its book-to-market ratio and its leverage ratio are included as control variables to remove the most significant effects from the regression model. While according to the authors the market value of a firm has a significant negative impact on its stock returns, the effect of BTM ratios is significantly positive. Additionally, market risk betas are added to the regression model, even though previous research on the investigated relation has shown no significant power of betas explaining returns (Galema, Plantinga & Scholtens, 2008; Mănescu, 2011). Consequently, the resulting cross-sectional regression model used in this paper looks as follows:

$$r_{i,t} = \gamma_0 + \gamma_1 * ESG_{i,t-X} + \gamma_2 * \beta_{i,t} + \gamma_3 * MV_{i,t-1} + \gamma_4 * BTM_{i,t-1} + \gamma_5 * lev_{i,t-1} + \varepsilon_{i,t} \quad (2)$$

where $r_{i,t}$ reflects monthly stock returns, $ESG_{i,t-X}$ the aggregated lagged Sustainalytics ESG score and $\varepsilon_{i,t}$ the error term. $\beta_{i,t}$, $MV_{i,t-1}$, $BTM_{i,t-1}$ and $lev_{i,t-1}$ represent all control variables in the model of which the last three are lagged by one month. The lag in the ESG variable is defined as *X*. This is due to the fact that the overall regression is performed three times using different lags ($ESG_{i,t-1}, ESG_{i,t-6}, ESG_{i,t-12}$) in order to test if investors might need longer time to price in ESG scores. The cross-section is regressed for each firm *i* = 1, … , 289 out of the original sample and repeated for every month *t* = 1, … , 123. Afterwards, time-series averages for all coefficients from the month-by-month cross-sectional regressions are calculated, representing the overall coefficient for every independent variable. These overall coefficients are then tested for significance using the following t-test:

$$t = \overline{\beta_k} \Big/ \frac{\sigma_{\widehat{\beta_k}}}{\sqrt{T}} \quad (3)$$

where $\overline{\beta_k}$ indicates the mean of a time-series average, $\sigma_{\widehat{\beta_k}}$ its standard deviation and T the sample size. The statistical significance is tested at 1%, 5% and 10% level.

Besides regressing the above model only on aggregated ESG scores, the impact of each pillar score on stock returns is investigated in a separate regression. Therefore, ESG_i in the above regression model is replaced by its respective pillar scores once a time. As a result, the overall effect of ESG scores on stock returns can be broken down into three partial effects, of which some can be more significant than others. Taking this knowledge into account for the later portfolio construction, financial performance might vary.

5.2 Pooled Regression on Stock-level

After having investigated the overall effect of ESG scores on financial stock performance using a cross-sectional regression, this section uses pooled regressions to analyze potential industry and country specific effects of ESG scores. The need to investigate these effects can be found in previous literature. Studies like Ziegler, Schröder and Rennings (2007) and Mănescu (2011) state a clear dependency between the ESG score and the industry a firm is operating in. Ignoring these industry and country specific effects could result in false inferences regarding effects of ESG integration into investment decisions. Since however that is the goal of this paper, a pooled regression approach is necessary to perform.

As already outlined in chapter 4, the overall sample analyzed in this paper includes 289 firms from 12 countries, covering 13 broad industries. However, some countries and industries are only represented by a few firms, making a cross-sectional regression for the investigation of potential industry and country specific effects inappropriate. A pooled regression bears the advantage of an increased sample size by treating the time-series dimension of a panel data set as an increased cross-section (Brooks, 2014, p. 527). By using this approach, the overall panel sample of 289 firms measured at 123 points in time can be treated as one big cross-sectional sample of 35,547 firms measured at only one point in time. The increased number of total cross-sections thus results in a higher number of observations per country and industry which consequently allows the investigation of potential industry and country specific effects with higher accuracy. Looking at the following pooled regression models, dummy variables representing countries and industries are used:

$$r_{i,t} = \beta_0 + \sum_{n=1}^{13} \left(\beta_n * ID_{i,n} + \beta_{n+13} * \left[ID_{i,n} * ESG_{i,t-1}\right]\right) + \sum_{j=27}^{29} \left(\beta_j * C_{i,j,t-1}\right) + \varepsilon_{i,t} \qquad (4)$$

$$r_{i,t} = \beta_0 + \sum_{n=1}^{12} \left(\beta_n * CD_{i,n} + \beta_{n+12} * \left[CD_{i,n} * ESG_{i,t-1}\right]\right) + \sum_{j=26}^{28} \left(\beta_j * C_{i,j,t-1}\right) + \varepsilon_{i,t} \qquad (5)$$

where $ID_{i,n}$ and $CD_{i,n}$ represent the industry respectively the country dummies that take on values of 1 if a firm operates in a respective industry or comes from a respective country and 0 otherwise. Furthermore, $C_{i,j,t-1}$ represents all control variables used in the cross-sectional regression model above. All remaining notations remain as before. As one can see, each dummy appears twice in both equations, once stand-alone and once as a product of ESG score and dummy variable. This is to clean the desired effects from general return patterns. While stand-alone dummies capture general return patterns in different countries and industries, the product of ESG score and dummy captures the desired clean country or industry specific ESG effects.

5.3 Time-series Regression on Portfolio-level

After the preceding sections have analyzed the relationship between ESG scores and financial performance on a stock-level, this section goes one step further into the direction of actual investment strategies by taking the investigated relationship on a portfolio-level. Therefore, a variety of different portfolios will be constructed, all having the similarity of taking social responsibility into account. The aim of this section is then to implement time-series regressions of portfolio returns that shed light on the research question whether there is a financial trade-off in socially responsible investments or not. In a first step, the next subsection describes the Fama and French four-factor model involved in the time-series regression in more detail.

5.3.1 Fama and French Four-Factor Model

As an extension to their cross-sectional approach in 1992, Fama and French (1993) further widened their asset pricing tests by using the time-series regression approach of Black, Jenson and Scholes (1972). In their original model, the time-series of monthly excess stock returns is regressed on both excess market returns and returns of zero-investment portfolios that are used as explanatory variables. By looking at the intercept of such a model, a simple return metric can be observed since according to Merton (1973), the intercept of a well-defined asset-pricing model should be indistinguishable from zero. Picking up this connection, monthly excess stock returns will be replaced by monthly excess returns of sustainable portfolios. Keeping the explanatory variables unchanged, the intercept of this paper's model can then be interpreted as the risk-adjusted return of a sustainable portfolio. Starting with their original model, Fama and French (1993) established a three-factor time-series model that looks as follows:

$$R_{i,t} - RF_t = \alpha_i + \beta_i * (RM_t - RF_t) + s_i * SMB_t + h_i * HML_t + \varepsilon_{i,t} \qquad (6)$$

where $R_{i,t} - RF_t$ represents an asset's excess return, $RM_t - RF_t$ the excess return of the market portfolio and SMB_t and HML_t the return of two zero-investment portfolios. Each zero-investment portfolio represents a common risk factor that is closely connected to a stock's return. SMB_t captures size effects of firms, HML_t captures their value effects. Looking at the construction of these two risk factor portfolios, all assets of an investment universe are first ranked into different sub groups. For the size factor, the median of market values separates the investment universe into two groups, small firms (S) and big firms (B). For the value factor, the 30_{th} and 70_{th} quantiles of the investment universe ranked according to book-to-market ratios are used to sort firms into three groups, low (L), medium (M) and high (H). In a second step, six different sub portfolios (SL, SM, SH, BL, BM, BH) are constructed by combining two rankings a time. The return of each sub portfolio is then calculated as the market value weighted return of all assets included:

$$XY = \sum_{i \in XY} \frac{MV_{i,t-1}}{\sum_{j \in XY} MV_{j,t-1}} * r_{i,t} \tag{7}$$

where XY illustrates the return of each respective sub portfolio. Finally, the overall return of both risk factor portfolios SMB_t and HML_t is calculated as follows:

$$SMB = \frac{SL + SM + SH}{3} - \frac{BL + BM + BH}{3} \tag{8}$$

$$HML = \frac{SH + BH}{2} - \frac{SL + BL}{2} \tag{9}$$

While the SMB_t (small minus big) risk factor portfolio represents the difference in returns between small and big firms, the HML_t (high minus low) portfolio is defined similarly using high and low valued firms. The main advantage using this approach of risk factor construction is that estimated coefficients in a time-series regression are comparable across portfolios that are used as dependent variables (Fama and French, 1993). This will be especially helpful when it later in this paper comes to the integration of ESG scores into existing investment strategies.

In order to get an asset pricing model that is even more precise than Fama and French’s (1993) three-factor model, this paper further adds Carhart’s (1997) momentum factor to its model. A momentum factor captures the risk of potential up- or downward trends in stock prices and is thus essential in describing stock returns. For its construction, the investment universe is first sorted by historical returns between t – 2 and t – 12. The resulting ranking is then split at the median in two groups, up (U) and down (D). Taking also the size-ranked groups small (S) and big (B) into account, the construction of this risk factor uses four different sub portfolios (US, UB, DS, DB) where the decreasing return sub portfolios are subtracted from the increasing ones to get to an overall UMD_t (up minus down) factor. The final four-factor model that will be used in a time-series regression to measure risk-adjusted returns of socially responsible portfolios now looks as follows:

$$ESG_t - RF_t = \alpha_i + \beta_i * (RM_t - RF_t) + s_i * SMB_t + h_i * HML_t + m_i * UMD_t + \varepsilon_{i,t} \quad (10)$$

where $ESG_t - RF_t$ represents the monthly excess return of a portfolio following a socially responsible investment strategy. In the next section of this paper, these ESG portfolios will now be constructed in line with different investment approaches.

5.3.2 Portfolio Construction

Even though in practice ESG scores generally do not build the basis of a modern investment strategy, most academic literature implicitly build their research framework upon this assumption (van Duuren, Plantinga & Scholtens, 2016). Therefore, this paper also first analyses the returns of pure ESG portfolios that are constructed solely on the basis of ESG scores. However, in order to implement a more realistic framework where ESG scores act as an add-on to existing investment strategies, an additional set of portfolios will be tested where ESG measures are subsequently integrated into existing investment strategies.

5.3.2.1 Pure ESG Portfolios

For the construction of pure ESG portfolios, three different methods are applied. While two methods are based on different zero-investment approaches, the third method represents a long-only strategy. Overall, by analyzing risk-adjusted returns of all three portfolios, the third supplementary hypothesis of this paper can be tested.

HML_{ESG} based on Fama and French

The first pure ESG portfolio is constructed based on the same methodology that Fama and French (1993) use to construct their HML risk factor (see section 5.3.1). However, instead of sorting the investment universe according to book-to-market ratios, it will be sorted according to its ESG scores. Using again the 30_{th} and 70_{th} quantiles of the ranked investment universe, firms are sorted into three groups, low ESG (L_{ESG}), medium ESG (M_{ESG}) and high ESG(H_{ESG}). Finally, the first pure ESG portfolio, HML_{ESG}, is then constructed like in equation 9. In this setup, both the high and low ESG sub portfolios have about the same weighted-average size.

HML_{ESG} based on Asness et al.

For the second pure ESG portfolio, a construction framework introduced by Asness, Ilmanen, Israel and Moskowitz (2015) is used. Therefore, the investment universe is again first ranked according to its ESG scores. Then, each rank is standardized by subtracting the mean ESG score from it and dividing the result by the standard deviation of all ESG scores. This setup assigns positive or negative weights to each asset, which in total add up to zero. In a next step, all weights are normalized so that all negative (positive) weights add up to -1 (1). In this framework, firms with a higher ESG score also get assigned a higher weight in the final portfolio.

Long-only based on EURO STOXX ESG Leaders 50

For the third ESG portfolio, a long-only investment strategy based on the EURO STOXX ESG Leaders 50 index is applied. This index provides an overview of global ESG leaders based on ESG indicators provided by Sustainalytics. By simply investing in this index, a pure ESG investment strategy is followed.

5.3.2.2 ESG Integration into existing Investment Strategies

Since in practice sustainability criteria rather act as an add-on to existing investment strategies, this subsection first constructs a range of zero-investment portfolios that solely follow their respective investment strategy. While all strategies are implemented using the same approach of standardized asset weights as in Asness et al. (2015), each investment strategy has a different style signal by which the investment universe is sorted. *Table 2* below shows all investment strategies and their respective style signals that are covered in this paper:

Table 2: Investment strategies

Each investment strategy is implemented based on the approach of zero-investment portfolios as in Asness et al. (2015). For every strategy, the investment universe is ranked according to its respective style factor. Then, according to "Description", high(low) ranked stocks are bought and low(high) ranked stocks are (short)sold. In total, the overall portfolio weight sums up to zero.

Strategy name	*Description*	*Style factor*
Size	Buy small and sell large firms	Current market capitalization
Value	Buy cheap and sell expensive stocks	Book-to-market ratio
Growth	Buy growing and sell declining firms	Three-year average EPS growth rate
Low Volatility	Buy less volatile and sell more volatile stocks	Standard deviation over the past 12 month
Momentum	Buy winners and sell losers	Historical return between t-12 and t-2
Quality	Buy high and sell low quality firms	Return on equity

After evaluating the financial stock performance of each investment strategy using the time-series framework explained above, ESG scores are added as a second ranking criteria to every strategy. Thus, a portfolio pays attention to compliance with both its strategy and ESG criteria. For the integration of ESG scores into a respective style portfolio, this paper follows a procedure based on Kaiser (2020). In order to obtain an overall style score that will be used to assign standardized weights, both the rank of a style signal together with the ESG rank of an asset are taken into account:

$$overall\ style\ score_i = 0.5 * rank(style\ signal_i) + 0.5 * rank(ESG) \quad (11)$$

Using this overall style score to assign standardized weights both lets the portfolio follow its respective investment strategy but also takes sustainability criteria into account. Thus, financial stock performance can be measured after the integration of ESG scores into existing investment strategies. In a last step, risk adjusted returns will be compared between a style portfolio without ESG integration and the same style portfolio with ESG integration. If both returns do not differ significantly from each other, the second sub-hypothesis of no financial trade-off in socially responsible investment strategies can be stated true.

5.3.3 Testing for equality of intercepts

In order to test if two risk adjusted returns do not differ significantly from each other (H_0: $\alpha_1 = \alpha_2$), hypotheses tests for equality of intercepts are applied. For this test, a test statistic is calculated using the difference between both risk-adjusted returns (intercepts) divided by the standard error of the difference between both intercepts:

$$Z = \frac{\alpha_2 - \alpha_1}{\sqrt{Var(\alpha_2 - \alpha_1)}} = \frac{\alpha_2 - \alpha_1}{\sqrt{Var(\alpha_1) + Var(\alpha_2) - 2 * Cov(\alpha_1, \alpha_2)}}$$

$$\approx \frac{\alpha_2 - \alpha_1}{\sqrt{(SE_{\alpha_1})^2 + (SE_{\alpha_2})^2}} \quad (12)$$

Since the calculation of the standard error of the difference between both intercepts requires their unknown covariance, this paper approximates the test statistic using both intercepts' standard errors obtained from the time-series regression (Clogg, Petkova & Haritou, 1995). If the test statistic is significant at either 1%, 5% or 10%, H_0 will be rejected, saying that both intercepts differ significantly from each other. If this is the case, one has to look which portfolio delivers higher risk-adjusted returns. If on the other side H_0 will not be rejected, the main hypothesis of no financial trade-off in socially responsible investments can be stated true, since financial performance does not significantly changes but sustainable performance clearly increases.

6 Results and Discussion

In this chapter, estimated results of both analyses on stock-level and portfolio-level are presented and discussed in order to shed light on all hypotheses developed earlier in this paper. Therefore, consistent with the methodological order of the last chapter, results of the cross-sectional regression as well as the pooled regression on stock-level are presented first to answer the sub-hypothesis dealing with the effect of ESG scores on financial stock performance of a firm. Once the relationship between ESG scores and stock performance is clear, focus is turned to the results of the time-series regression on portfolio level. By evaluating financial stock performance of different portfolios using a four-factor asset-pricing model, the aim is to confirm the remaining two sub-hypotheses and finally the main hypothesis of no financial trade-off in socially responsible investments.

6.1 Cross-sectional Regression

Starting with the cross-sectional regression model using aggregated ESG scores, *Table 3* below illustrates the estimated results for three different time lags in ESG scores. By regressing the same model three times with differently lagged ESG scores, this paper aims to investigate how long investors need to price in the announcement of an updated ESG score. This analysis can be seen as a preliminary investigation since its outcome will form the basis for all upcoming regressions. Looking into the results, one can see that all time lags result in an almost similar ESG estimate that is significant at least at 5% level. Thus, ESG scores are already priced in immediately after their announcement and the effect does not change over time. This finding is consistent with previous literature like Galema, Plantinga and Scholtens (2008) and Mănescu (2011), who also use a time lag of only one month for their regression models. Therefore, in the further course of this paper, all regression models will be performed using only one month lag in ESG scores. Again, a lag of at least one month is necessary since we are looking at monthly returns that cannot reflect ESG scores in the same moment of their announcement.

Turning the focus on the outcome of the cross-sectional regression, one can see that besides ESG scores also market capitalization and leverage of a firm significantly help to explain monthly stock returns, which is in line with the original findings from Fama and French (1992). However, for this paper's sample, book-to-market ratios do not significantly help to explain the cross-section of returns. Furthermore, as already found in previous literature, market risk betas also have no significant power in explaining the cross section of stock returns. Most importantly, aggregated ESG scores have a significant negative relation to monthly stock returns. This states that firms with high ESG scores tend to generate lower returns than firms with low ESG scores, which rejects this paper's first sub-hypothesis. A possible explanation for this result can be found in Derwall and Verwijmeren (2007), where the authors justify negative returns of highly rated stocks with a too low compensation for risk connected to social responsibility.

Table 3: Cross-sectional regression using different ESG time lags

This table summarizes the time-series average of monthly cross-sectional regressions performed between September 2009 and November 2019. While monthly stock returns serve as dependent variable, lagged ESG scores together with a set of control variables serve as explanatory variables. The first value for each variable represents its estimated coefficient while the second value (in parentheses) shows the t-statistic.

	1 month ESG lag	6 months ESG lag	12 months ESG lag
Intercept	0.0200*** (4.6013)	0.0186*** (4.2212)	0.0187*** (4.1577)
ESG_{t-X}	-0.0002*** (-2.9350)	-0.0002** (-2.5770)	-0.0001** (-2.5508)
$Beta_t$	-0.0014 (-0.4179)	-0.0015 (-0.4237)	-0.0016 (-0.4465)
MV_{t-1}	-0.0000** (-2.4828)	-0.0000** (-2.4020)	-0.0000** (-1.9255)
BTM_{t-1}	-0.0001 (-0.0255)	0.0005 (0.2358)	0.0002 (0.0850)
lev_{t-1}	-0.0009*** (-2.9350)	-0.0008** (-2.5770)	-0.0008** (-1.9133)

Note: * p < 0.1; ** p < 0.05; *** p < 0.01

Taking a closer look at the three pillars of ESG, namely environmental practices, social relations and corporate governance, the above finding of a negative relation to monthly stock returns can also be found in each pillar score. As reported in *Table 4*, each pillar score has a negative coefficient that is significant at least at 5% level. However, the size of the coefficients differs. While the social pillar has the highest negative impact on stock returns, the environmental and governance pillars follow in decreasing order. This result states that investors attach varying degrees of importance to the individual pillars of ESG. Even if all components have an overall negative impact on stock returns, investors tend to price social aspects more heavily than for example topics that are related to the corporate governance of a firm. This finding is largely consistent with most preceding literature related to ESG scores. While social and environmental performance seem to be most important for socially responsible investors, corporate governance scores often only explain returns insignificantly.

Table 4: Cross-sectional regression using different ESG pillar scores

This table summarizes the time-series average of monthly cross-sectional regressions performed between September 2009 and November 2019. While monthly stock returns serve as dependent variable, environmental, social and governance pillar scores together with a set of control variables serve as explanatory variables. The first value for each variable represents its estimated coefficient while the second value (in parentheses) shows the t-statistic.

	Environmental pillar	Social pillar	Governance pillar
Intercept	0.0164*** (4.7359)	0.0195*** (4.6787)	0.0164*** (3.9562)
E_{t-1} / S_{t-1} / G_{t-1}	-0.0001*** (-2.8455)	-0.0002*** (-3.2302)	-0.0000** (-1.9094)
$Beta_t$	-0.0016 (-0.4782)	-0.0013 (-0.4039)	-0.0015 (-0.4623)
MV_{t-1}	-0.0000*** (-2.6342)	-0.0000*** (-2.5521)	-0.0000*** (-2.9365)
BTM_{t-1}	-0.0003 (-0.1428)	0.0002 (0.0710)	-0.0003 (-0.1552)
lev_{t-1}	-0.0008** (-2.0694)	-0.0009** (-2.1985)	-0.0009** (-2.1906)

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Summarizing all results of the cross-sectional regressions above, it can be stated that a high degree of social responsibility tends to lower a firm's financial stock performance. However, magnitudes differ with each ESG pillar and are generally rather small.

6.2 Pooled Regression

As already explained in chapter 5, pooled regressions are performed to identify potential industry and country specific patterns in ESG scores. However, it can also be used to verify the results of the above cross-sectional regression. By performing a pooled regression without dummy variables, the regression model is identical to the one performed in section 5.1, making results comparable. As one can see in *Table 5*, a pooled regression without dummy variables leads to a similar result as the time-series average of cross-sectional regressions reported in *Table 3*. Most importantly, the ESG coefficient is still significantly negative, even if its magnitude is smaller. Interestingly, the explanatory power of all control variables further increases in a pooled regression where also market risk betas become significant at 1% level.

Table 5: Pooled regression to verify cross-sectional results

This table summarizes the results of a pooled regression where the panel data set is treated as one large cross-section. While monthly stock returns serve as dependent variable, lagged ESG scores together with a set of control variables serve as explanatory variables. The first value for each variable represents its estimated coefficient while the second value (in parentheses) shows the t-statistic.

	Intercept	ESG_{t-1}	$Beta_t$	MV_{t-1}	BTM_{t-1}	lev_{t-1}
Coefficient	0.014***	-0.001*	-0.004***	-0.000***	0.006***	-0.001***
(t-statistic)	(4.7537)	(-1.9195)	(-2.9876)	(-4.6666)	(5.1242)	(-5.1350)

Note: * p < 0.1; ** p < 0.05; *** p < 0.01

Now, that the results of the cross-sectional regression have been validated, the next step is the evaluation of country and industry specific ESG patterns. Therefore, *Table 6* reports the results of a pooled regression using industry dummies and *Table 7* shows results using country dummies.

Table 6: Pooled regression measuring industry-specific effects

This table summarizes the results of a pooled regression using industry dummies for a period between September 2009 and November 2019. While monthly stock returns serve as dependent variable, industry dummies are integrated into the regression model to represent industry-specific effects. Each dummy is used twice, once as the product of dummy and ESG score and once stand-alone. This table represents the coefficients of the product dummy. The whole table can be found in the appendix.

	Coefficient	t-statistic
Intercept	-0.0502**	-2.1052
Beta $_{t}$	-0.0027***	-4.1332
MV $_{t-1}$	-0.0001***	-4.1332
BTM $_{t-1}$	0.0101***	8.2336
lev $_{t-1}$	-0.0002***	-4.3114
Utilities	0.0002	1.6364
Financial	0.0007	0.7873
Real Estate	0.0005	0.3159
Transportation	-0.0003	-0.1487
Automobile	-0.0009	-0.4227
IT	-0.0005***	-4.4734
Construction	0.0008**	2.2287
Food	-0.0002	-0.5758
Machinery	-0.0001	-0.5906
Pharma & Chemicals	-0.0002*	-1.7710
Healthcare	0.0003	0.1578
Textile & Apparel	-0.0008	-0.3692

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Starting with the t-statistics of control variables, one can observe that all four are significant at 1% level in both regressions. While market risk betas, market values and leverage ratios of a firm have a negative effect on its financial stock returns, the effect of book-to-market ratios is positive. These results are in line with the original findings in Fama and French (1992). Moving to the industry specific effects in *Table 6*, one can clearly see that coefficients differ between industries. More precisely, some industries even have coefficients with different signs, which confirms that the relation between ESG scores and financial stock returns varies depending on the industry a firm is operating in. However, the majority of industries only deliver coefficients that are below a 10% significance level, making these results hard to interpret. Still, some industries show significant industry specific ESG effects. While in industries like Information Technologies (IT) or Pharma and Chemicals ESG scores have a significant negative relation to stock returns, results for the Construction industry are different. Here, ESG scores have a significant positive relation to stock returns, confirming the occurrence of industry specific ESG effects. Turning the focus to country specific effects in *Table 7*, similar results can be found.

Table 7: Pooled regression measuring country-specific effects

This table summarizes the results of a pooled regression using industry dummies for a period between September 2009 and November 2019. While monthly stock returns serve as dependent variable, country dummies are integrated into the regression model to represent country-specific effects. Each dummy is used twice, once as the product of dummy and ESG score and once stand-alone. This table represents the coefficients of the product dummy. The whole table can be found in the appendix.

	Coefficient	t-statistic
Intercept	0.0050	1.1950
$Beta_t$	-0.0036***	-2.9206
MV_{t-1}	-0.0000***	-4.5576
BTM_{t-1}	0.0071***	6.0993
lev_{t-1}	-0.0009***	-4.8319
Austria	0.0003	1.3664
Belgium	0.0001	0.6483
Finland	-0.0002	-0.8106
France	-0.0001	-1.2377
Germany	-0.0002**	-2.0451
Ireland	0.0001	0.4122
Italy	0.0002	0.9984
Luxembourg	-0.0003	-0.7668
Netherlands	-0.0002	-1.4577
Spain	0.0000	0.2650
Switzerland	0.0036	1.2308
United Kingdom	-0.0010**	-2.3601

Note: * p < 0.1; ** p < 0.05; *** p < 0.01

While some countries show positive coefficients, others show negative ones. However, the majority of countries again deliver insignificant coefficients, resulting in a restriction on the ability of interpretation. Two exceptions are Germany and the United Kingdom, where a significant negative ESG effect on monthly stock returns can be found. Overall, even if results are hard to interpret due to a lack of significance, some industry and country specific ESG effects can still clearly be confirmed. Therefore, the result of an overall negative relationship between ESG scores and financial stock performance, obtained from the cross-sectional regression above, has to be taken with caution. By adjusting the analyzed sample regarding some specific industries or countries, results might most likely change. However, these adjustments are not part of this paper's analysis. In the further course of this chapter, focus will be put on whether portfolios constructed according to ESG criteria generate significant positive risk-adjusted returns even though the relation obtained in the cross-sectional regression is negative. Should this be the case, the main hypothesis of no financial trade-off in socially responsible investments can finally be confirmed.

6.3 Time-series Regression

6.3.1 Pure ESG Portfolios

In a first step, empirical results of time-series regressions using a four-factor asset-pricing model are presented and discussed for portfolios that are constructed solely on the basis of ESG scores. While the first two portfolios represent zero-investment strategies where assets with high ESG scores are bought and those with low scores are (short)sold, the third portfolio represents a long-only strategy where investments are made in the EURO STOXX ESG Leaders 50 index. As reported in *Table 8* below, all three portfolios show positive intercepts alpha that are significant at 1% level. Consistent with the theory of asset-pricing models, these intercepts can be interpreted as positive risk-adjusted returns, proving this paper's third supplementary sub-hypothesis that pure ESG portfolios still come with an appropriate financial performance.

Table 8: Time-series regression for pure ESG portfolios

This table summarizes the results of a time-series regression performed between January 2014 and November 2019. Monthly excess portfolio returns are explained using a Fama and French four-factor model including Carhart's momentum factor. The first value for each variable represents its estimated coefficient while the second value (in parentheses) shows the t-statistic. While, Portfolio one (I) and two (II) are zero-investment portfolios, portfolio three (III) is a long-only portfolio investing in the EURO STOXX ESG leaders 50 index.

	(I) HML Fama and French	(II) HML Asness et al.	(III) Long-only EURO STOXX ESG leaders 50
Intercept	0.0135*** (3.1699)	0.0178*** (5.6700)	0.0219*** (5.6588)
Rm – Rf	-0.6499*** (-5.2575)	0.1298 (1.4290)	0.8971*** (7.7445)
SMB	0.4072 (1.4853)	-0.5243** (-2.6027)	-0.8459*** (-3.3426)
HML	0.6257* (2.7135)	0.5276*** (3.1137)	0.0529 (0.2489)
UMD	0.22604 (1.135)	0.1013 (0.6919)	-0.0031 (0.0170)

Note: * p < 0.1; ** p < 0.05; *** p < 0.01

Furthermore, some interesting inferences can be drawn from the common risk factors of the model. First, excess returns of the second and third ESG portfolio have a significant negative relation to Fama and French's SMB portfolio. This result allows the inference that small firms most likely have lower ESG scores compared to large firms. Second, both zero-investment ESG portfolios are significantly positively correlated with the HML portfolio, indicating that firms with higher values also seem to have higher ESG scores. Finally, looking at Carhart's momentum factor, one can observe that both zero-investment portfolios deliver positive, yet insignificant coefficients. Assuming the results were significant, this scenario would allow the inference that firms with higher ESG scores have also generated higher returns in the past. However, since results are insignificant, this inference has to be taken with caution.

6.3.2 ESG Integration into existing Investment Strategies

Finally, this section presents risk-adjusted returns from existing investment strategies measured once with and once without ESG integration. Using the same four-factor asset-pricing model as for the pure ESG portfolios, *Table 9* is divided into six different panels where each panel represents one investment strategy.

Table 9: Time-series regression for existing investment strategies

This table summarizes the results of a time-series regression performed between January 2014 and November 2019. Monthly excess portfolio returns are explained using a Fama and French four-factor model including Carhart's momentum factor. The first value for each variable represents its estimated coefficient while the second value (in parentheses) shows the t-statistic. While, for each panel the first (I) portfolio represents a strategy without ESG incorporation, the second portfolio (II) represents the same strategy with ESG incorporation.

Panel A: Size

	Intercept	Rm – Rf	SMB	HML	UMD
(I)	0.0195***	0.0539	-0.7103***	0.1555	0.1705
	(5.5850)	(0.5332)	(-3.1656)	(0.8238)	(1.0460)
(II)	0.0180***	0.1164	-0.5188**	0.4715***	0.1030
	(5.7658)	(1.2888)	(-2.5906)	(2.7989)	(0.7079)

Panel B: Value

	Intercept	Rm – Rf	SMB	HML	UMD
(I)	0.0079*	0.2986**	0.0539	1.5154***	-0.5027***
	(1.9705)	(0.2986)	(0.2103)	(7.0176)	(-2.6953)
(II)	0.0178***	0.1174	-0.5022**	0.4946***	0.0837
	(5.7475)	(1.3063)	(-2.5204)	(2.9505)	(0.5785)

Panel C: Growth

	Intercept	Rm – Rf	SMB	HML	UMD
(I)	0.0212***	0.1746	0.0899	-0.1235	-0.0706
	(2.9488)	(0.8395)	(0.1949)	(-0.3182)	(-0.2108)
(II)	0.0181***	0.1124	-0.5101**	0.4399**	0.1079
	(5.8270)	(1.2471)	(-2.5519)	(2.6162)	(0.7428)

Panel D: Low Volatility

	Intercept	Rm – Rf	SMB	HML	UMD
(I)	0.0238*** (5.7553)	0.2772** (2.3071)	0.0099 (0.0391)	0.7332*** (3.2721)	0.3301* (-1.7056)
(II)	0.0182*** (5.8597)	0.1170 (1.2973)	-0.5093** (2.5462)	0.4840*** (2.8768)	0.0884 (0.6082)

Panel E: Momentum

	Intercept	Rm – Rf	SMB	HML	UMD
(I)	0.0119*** (3.1038)	0.0270 (0.24409)	0.1779 (0.7244)	-0.2442 (-1.1820)	1.1804*** (6.6149)
(II)	0.0177*** (5.6913)	0.1153 (1.2780)	-0.5085** (-2.5423)	0.4650*** (2.7638)	0.1357 (0.9339)

Panel F: Quality

	Intercept	Rm – Rf	SMB	HML	UMD
(I)	0.0260*** (6.3267)	-0.0204 (-0.1706)	0.3149 (1.1903)	-0.6748*** (-3.0322)	0.2334 (1.2142)
(II)	0.0182*** (5.8298)	0.1133 (1.2513)	-0.5179** (-2.5787)	0.4500*** (2.6639)	0.1028 (0.7042)

Note: * p < 0.1; ** p < 0.05; *** p < 0.01

The first important finding that can be made in all panels is that each investment strategy has a positive and significant intercept, meaning that they all have earned positive risk-adjusted returns when ESG criteria are not taken into account. While growth, quality, momentum and low volatility strategies have earned average monthly risk-adjusted returns of around two percent, the remaining two strategies have returned between 0.8% and 1.2% on average per month. Interestingly, the low volatility strategy earned the second highest return, even though its aim is to encounter the lowest risk. Turning focus to those strategy portfolios where ESG scores are additionally taken into account, all intercepts remain positively significant. While value and momentum styles even benefit from the incorporation of ESG scores by earning slightly higher returns afterwards, the remaining strategies experience some slight loss of return. Furthermore, as one can see in *Figure 1* (appendix), also the volatility of returns improved after the incorporation of ESG scores for five out of six strategies, only leaving the size strategy unchanged. In order to test whether differences in risk-adjusted returns are significant or not, tests for equality of intercepts are performed and results are presented in *Table 10* below.

Table 10: Hypothesis tests for equality of intercepts

This table summarizes the results of different tests for equality of intercepts. Intercept coefficients as well as their respective standard errors are used to calculate the test-statistic as described in equation 12. While, for each panel the first (I) portfolio represents a strategy without ESG incorporation, the second portfolio (II) represents the same strategy with ESG incorporation.

Panel A: Size

	Intercept	SE	test-statistic (Z)
(I)	0.0195***	0.00349	-0.3273
(II)	0.0180***	0.00311	

Panel B: Value

	Intercept	SE	test-statistic (Z)
(I)	0.0079*	0.00399	1.9676**
(II)	0.0178***	0.00310	

Panel C: Growth

	Intercept	SE	test-statistic (Z)
(I)	0.0212***	0.00717	-0.3888
(II)	0.0181***	0.00311	

Panel D: Low Volatility

	Intercept	SE	test-statistic (Z)
(I)	0.0238***	0.00414	-1.0844
(II)	0.0182***	0.00311	

Panel E: Momentum

	Intercept	SE	test-statistic (Z)
(I)	0.0119***	0.00382	1.1872
(II)	0.0177***	0.00311	

Panel F: Quality

	Intercept	SE	test-statistic (Z)
(I)	0.0260***	0.00412	-1.5143*
(II)	0.0182***	0.00312	

Note: * p < 0.1; ** p < 0.05; *** p < 0.01

As one can see, only two out of six investment strategies report a significant test-statistic that indicates a significant change in risk-adjusted returns after the incorporation of ESG scores. While for panels A, C, D and E the respective test-statistic lies below the 10% level threshold, only panels B and F deliver a significant test-statistic. First, the value strategy in panel B increased its risk-adjusted return from 0.79% to 1.78% at 5% significance level. Second, returns for the quality strategy in panel F decreased from 2.60% to 1.82% at 10% significance level. Thus, it can be stated that overall, after the incorporation of ESG scores, only one strategy has significantly lost returns while the remaining five strategies either did not have significant changes or even yield a significant increase in risk-adjusted returns. To summarize the findings of all time-series regressions, this paper comes to the result that in most cases ESG scores can be used as a style factor to generate both financial and sustainable performance. While all pure ESG portfolios deliver significant positive risk-adjusted returns, the incorporation of ESG scores into existing investment strategies only comes with a financial trade-off once. In the remaining five cases, each strategy improves its sustainable performance without significantly reducing financial performance.

7 Conclusion

As already mentioned in the introduction to this paper, the present empirical study investigates the financial impact of socially responsible behavior on existing investment processes and decisions in Central Europe. For this purpose, the main hypothesis of no financial trade-off in socially responsible investments is formulated and subsequently analyzed using a data set covering the STOXX 600 index between September 2009 and November 2019. While financial performance of an investment is expressed by its monthly risk-adjusted returns, sustainable performance is expressed by ESG scores. In order to investigate the relation between both measures and thus shed light on this paper's research question, two methodological approaches are carried out. First, cross-sectional regressions of stock returns are applied as in Fama and French (1992). Using aggregated ESG scores as explanatory variables, this regression approach aims to capture the general relation between ESG scores and financial performance on a stock-level. For a deeper knowledge of this relationship, additional cross-sectional regressions using single pillar scores are performed to understand the individual effects of environmental, social and corporate governance aspects. Furthermore, pooled regressions are used to analyze potential industry and country specific patterns in ESG. Moving from a stock-level relation more into the direction of actual investment decisions, focus is then turned to a portfolio-level. Therefore, portfolios are constructed using ESG scores both as the basis of an investment strategy as well as an add-on to existing strategies. By implementing time-series regressions using a Fama and French four-factor asset pricing model, risk-adjusted returns for all portfolios are obtained. Finally, to investigate potential financial trade-offs due to an ESG incorporation into existing investment strategies, risk-adjusted returns are compared using hypothesis tests for equality of intercepts.

Looking at the main findings from the cross-sectional regression approach on stock-level, results indicate a significant negative relation between ESG scores and financial stock performance. While the relation also stays significantly negative for all pillar scores, it can be conducted that firms with lower ESG scores outperform their higher rated peers. Rejecting the first sub-hypothesis, these findings are largely consistent with preceding theoretical literature. A firm that moves towards a sustainable business minimizes its costs of conflicts and maximizes shareholder value by avoiding potential conflicts with society, however its new sustainability also comes at costs. The negative relation between ESG scores and financial stock performance can therefore most likely be explained by ESG costs exceeding a profitable level. However, this is not the case for every industry. Results from pooled regressions controlling for industry and country specific effects show that a careful selection of industries and countries can also result in an overall positive relationship. Thus, investment strategies do not necessarily suffer from the observed negative relation. This statement is also confirmed by the findings from the time-series regression approach on portfolio-level. Results show that regardless of whether ESG is treated as the basis of an investment strategy or as an add-on to existing strategies, risk-adjusted returns for all portfolios remain significantly positive. Furthermore, looking into the results of two portfolios following the same investment strategy, once with and once without ESG integration, hypotheses tests for equality of intercepts proof this paper's main hypothesis of no financial trade-off in socially responsible investments. While only one strategy has significantly lost returns when ESG is additionally taken into account, all remaining strategies either did not have significant changes or even yield a significant increase in risk-adjusted returns. Thus, a financial trade-off was only observable once while in the remaining five cases, each strategy improved its sustainable performance without significantly reducing financial performance.

Overall, empirical results provide evidence that ESG scores can be used as a style factor to generate both financial and sustainable performance without forcing a trade-off between both.

References

Allen, F., Carletti, E., & Marquez, R. (2007). Stakeholder capitalism, corporate governance and firm value, *working paper*, University of Pennsylvania

Asgharian, H. (2004). A Comparative Analysis of the Ability of Mimicking Portfolios in Representing the Background Factors, *working paper*, Lund University

Asness, C. S., Moskowitz, T. J., & Pedersen, L. H. (2013). Value and momentum everywhere, *Journal of Finance*, vol. 68, no. 3, pp. 929–985

Asness, C. S., Ilmanen, A., Israel, R., & Moskowitz, T. J. (2015). Investing with style, *Journal of Investment Management*, vol. 13, no. 1, pp. 27-63

Bello, Z. (2005). Socially responsible investing and portfolio diversification, *Journal of Financial Research*, vol. 28, no. 1, pp. 41-57

Belu, C. (2009). Ranking corporations in terms of sustainable and socially responsible practices, *Journal of Sustainable Development*, vol. 17, no. 4, pp. 257-268

Berry, T., & Junkus, J. (2013). Socially Responsible Investing: An Investor Perspective, *Journal of Business Ethics*, vol. 112, no. 4, pp. 707-720

Black, F., Jensen, M. C., & Scholes, M. (1972). The Capital Asset Pricing Model: Some Empirical Tests, in M. C. Jensen (eds), *Studies in the Theory of Capital Markets*, New York: Praeger

Brooks, C. (2014). Introductory Econometrics for Finance, 3rd edn, Cambridge: Cambridge University Press

Carhart, M. M. (1997). On Persistence in Mutual Fund Performances, *Journal of Finance*, vol. 52, no. 1, pp. 57-82

Charoenrook, A., & Conrad, J. (2005). Identifying risk-based factors, *discussion paper*, AFA 2006 Boston Meetings

Clogg, C. C., Petkova, E., & Haritou, A. (1995). Statistical Methods for Comparing Regression Coefficients Between Models, *American Journal of Sociology*, vol. 100, no. 5, pp. 1261-1293

Cohen, M. A., Fenn, S. A., & Konar, S. (1997). Environmental and financial performance: are they related?, *working paper*, Vanderbilt University

Dam, L. (2008). Corporate Social Responsibility and Financial Markets, *PhD thesis*, University of Groningen, Available online: https://www.rug.nl/research/portal/files/10346481/thesis.pdf [Accessed 18 May 2020]

Derwall, J., & Verwijmeren, P. (2007). Corporate Governance and the Cost of Equity Capital: Evidence From GMI's Governance Rating, *ECCE Research Note 06-01*, European Centre for Corporate Engagement

Derwall, J., Guenster, N., Bauer, R., & Koedijk, K. (2005). The eco-efficiency premium puzzle, *Financial Analyst Journal*, vol. 61, no. 2, pp. 51-63

European Central Bank (2020). *Statistical Data Warehouse – Real 3-month Euribor*, Available online: http://sdw.ecb.europa.eu/browseSelection.do;jsessionid=37E95AEE97F1BC82195E6E160377DBCB?type=series&q=euribor+rate&node=SEARCHRESULTS&ec=&oc=&rc=&cv=&pb=&dc=&df= [Accessed 27 April 2020]

Fama, E. F., & French, K. R. (1992). The Cross-Section of Expected Stock Returns, *Journal of Finance*, vol. 47, no. 2, pp. 427-465

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics*, vol. 33, no. 1, pp. 3-56

Fama, E. F., & French, K. R. (2007). Disagreement, tastes and asset prices, *Journal of Financial Economics*, vol. 83, no. 3, pp. 667-689

Fama, E. F., & MacBeth, J. (1973). Risk, Return and Equilibrium: Empirical Tests, *Journal of Political Economy*, vol. 81, no. 3, pp. 607-636

French, K. R. (2020). *Kenneth R. French Data Library*, Available online: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html [Accessed 27 April 2020]

Friede, G., Busch, T., & Bassen, A. (2015). ESG and Financial Performance: Aggregated Evidence from more than 2000 Empirical Studies, *Journal of Sustainable Finance and Investment*, vol. 5, no. 4, pp. 210-233

Galema, R., Plantinga, A., & Scholtens, B. (2008). The stocks at stake: Return and risk in socially responsible investment, *Journal of Banking and Finance*, vol. 32, no. 12, pp. 2646-2654

Giese, G., Lee, L., Melas, D., Nagy, Z., & Nishikawa, L. (2019). Foundations of ESG Investing: How ESG Affects Equity Valuation, Risk, and Performance, *Journal of Portfolio Management*, vol. 45, no. 5, pp. 69-83

Goldreyer, E. F., & Diltz, J. D. (1999). The performance of socially responsible mutual funds: incorporating sociopolitical information in portfolio selection, *Managerial Finance*, vol. 25, no. 1, pp. 23-36

Hamilton, J. (1995). Pollution as news: Media and stock market reactions to the toxics release inventory data, *Journal of Environmental Economics and Management*, vol. 28, no. 1, pp. 98-113

Hamilton, S., Joe, H., & Statman, M. (1993). Doing well while doing good? The investment performance of socially responsible mutual funds, *Financial Analysts Journal*, vol. 49, no. 6, pp. 62-66

Heal, G. (2005). Corporate social responsibility: An economic and financial framework, *The Geneva Papers on Risk and Insurance – Issues and Practice*, vol. 30, no. 3, pp. 387-409

Heinkel, R., Krause, A., & Zechner, J. (2001). The effect of green investment on corporate behavior, *Journal of Financial and Quantitative Analysis*, vol. 36, no. 4, pp. 431-449

Hong, H., & Kacperczyk, M. (2005). The price of sin: The effect of social norms of markets, *Journal of Financial Economics*, vol. 93, no. 1, pp. 15-36

Jensen, M. (2004). Agency costs of overvalued equity and the current state of corporate finance, *Journal of European Financial Management*, vol. 10, no. 4, pp. 549-565

Kaiser, L. (2020). ESG integration: value, growth and momentum, *Journal of Asset Management*, vol. 21, no. 1, pp. 32-51

Kempf, A., & Osthoff, P. (2007). The Effect of Socially Responsible Investing on Portfolio Performance, *Journal of European Financial Management*, vol. 13, no. 5, pp. 908-922

Klassen, R., & McLaughlin, C. (1996). The impact of environmental management on firm performance, *Journal of Management Science*, vol. 42, no. 8, pp. 1199-1214

Konar, S., & Cohen, M. A. (1997): Information as regulation: The effect of community right to know laws on toxic emissions, *Journal of Environmental Economics and Management*, vol. 32, no. 1, pp. 109-124

Konar, S., & Cohen, M. A. (2001). Does the market value environmental performance?, *Review of Economics and Statistics*, vol. 83, no. 2, pp. 281-289

Mănescu, C. (2011). Stock Returns in relation to Environmental, Social and Governance Performance: Mispricing or Compensation for risk?, *Journal of Sustainable Development*, vol. 19, no. 2, pp. 95-118

McWilliams, A., Siegel, D., & Teoh, S. H. (1999). Issues in the Use of the Event Study Methodology: A Critical Analysis of Corporate Social Responsibility Studies, *Organizational Research Methods*, vol. 2, no. 4, pp. 340-365

Melas, D., Nagy, Z., & Kulkarni, P. (2006). *Factor investing and ESG integration*, Available online: https://www.msci.com/www/research-paper/factor-investing-and-esg/0528147089 [Accessed 09 April 2020]

Merton, R. C. (1973). An Intertemporal Capital Asset Pricing Model, *Econometrica*, vol. 41, no. 5, pp. 867-887

Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information, *Journal of Finance*, vol. 42, no. 3, pp. 483–510

Morgan Stanley Institute for Sustainable Investing (2017). *Sustainable Signals - New Data from the Individual Investor*, Available on: https://www.morganstanley.com/pub/content/dam/msdotcom/ideas/sustainable-signals/pdf/Sustainable_Signals_Whitepaper.pdf [Accessed 19 April 2020]

Renneboog, L., Ter Horst, J., & Zhang, C. (2008a). Socially responsible investments: Institutional aspects, performance, and investor behavior, *Journal of Banking and Finance*, vol. 32, no. 9, pp. 1723-1742

Renneboog, L., Ter Horst, J., & Zhang, C. (2008b). The price of ethics and stakeholder governance: The performance of socially responsible mutual funds, *Journal of Corporate Finance*, vol. 14, no. 3, pp. 302-322

Scholtens, B., & Zhou, Y. (2008). Stakeholder relations and financial performance, *Journal of Sustainable Development*, vol. 16, no. 3, pp. 213-232

Sharpe, W. F. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk, *Journal of Finance*, vol. 19, no. 3, pp. 425-442

Statman, M. (2000). Socially Responsible Mutual Funds, *Financial Analysts Journal*, vol. 56, no. 3, pp. 30-39

Statman, M., & Glushkov, D. (2009). The Wages of Social Responsibility, *Financial Analysts Journal*, vol. 65, no. 4, pp. 33-46

US SIF Foundation (2019). *Sustainable and Impact Investing - Overview*, Available on: https://www.ussif.org/files/2018%20Infographic%20overview%20(1).pdf [Accessed 14 April 2020]

van Duuren, E., Plantinga, A., & Scholtens, B. (2016). ESG Integration and the Investment Management Process: Fundamental Investing Reinvented, *Journal of Business Ethics*, vol. 138, no. 3, pp. 525-533

Ziegler, A., Schröder, M., & Rennings, K. (2007). The effect of environmental and social performance on the stock performance of European corporations, *Journal of Environmental and Resource Economics*, vol. 37, no. 4, pp. 661-680

Appendix

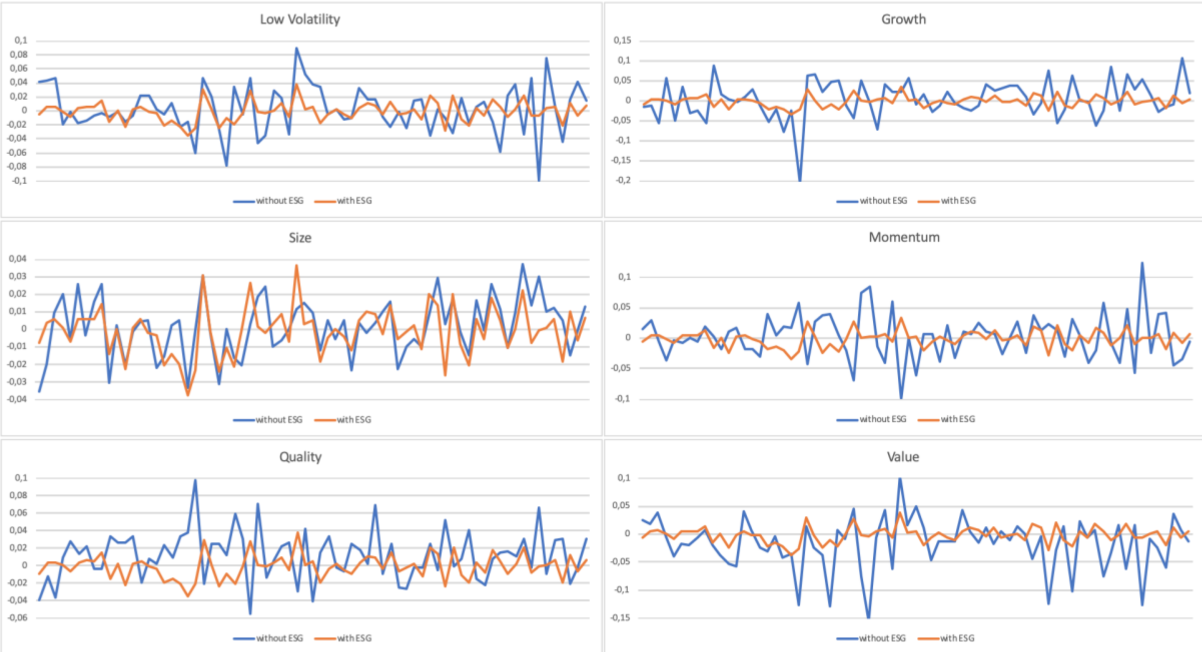

Figure 1: Volatility of strategy portfolios with and without ESG integration