LUND
UNIVERSITY

ESG Portfolios in Different Markets

Investigating the Relationship Between ESG Performance and Financial Performance

Author:
Ludwig Wolff

Supervisor: Hossein Asgharian

Department of Economics
NEKP01 – Master Essay II

August 2022

Abstract

By applying one of the largest datasets on ESG ratings to date with around 8000 companies included during the sample period between 2006-2021. This paper investigates the increasingly popular link between firms' social and financial performance and the potential abnormal returns to be found using ESG investment strategies. To examine this relationship, a series of portfolios are constructed based on firms with high and low ESG ratings as well as portfolios taking the difference between the high and the low portfolios to measure abnormal returns between the two. To check the results for robustness, additional portfolios are implemented based on variables such as portfolio size, subperiods, weighting method, and finally, firm selection criteria where a best-in-class approach is applied. The portfolios are then evaluated using the Carhart (1997) four-factor model. The result in this paper indicates that difference portfolios fail to find any consistent abnormal performance between investing in portfolios containing high ESG-rated firms and low ESG-rated firms across the three regions for the entire period. When looking at the individual high and low-rated portfolios, the low-rated portfolios frequently outperform their high counterpart, and when splitting the sample period in two and applying different robustness tests, some patterns of significant abnormal returns are found. The results also visibly show higher abnormal returns for the high (low) portfolios using equally weighted stocks than weighting based on market capitalization. The overall presence but inconsistency of the significant alphas in the portfolios suggest investors to be cautious but curious in the attempts of investment strategies based on ESG portfolios.

Keywords: ESG Portfolios, Abnormal Returns, Carhart Four-Factor Model, U.S, Europe, Emerging markets

Contents

1. Introduction

There is an old Swedish proverb that a dear child has many names, meaning something valuable is known by several names. This has undoubtedly been the case for the importance of firms' social performances in investing during the last decades, where the popularity has exploded exponentially, going under a considerable number of interchangeable terms during this time, such as corporate social responsibility (CSR), socially responsible investments (SRI), impact investing, and sustainable investing, to name a few. Nowadays, the more popular phrase for this umbrella term is ESG which stands for environmental, social, and governance.

The relationship between a firm's social and financial performance has been discussed for decades; where one of the first instances goes back to the early fifties when Bowen (1953), in his book, pondered about the long-run merge of interest between business and society and coined the term CSR, Corporate Social Responsibility. A concept that aimed to make a business accountable regarding sustainability. Around half a century later, the nowadays mainstream term, ESG, was introduced (UNEP FI, 2005), which has become a tool to measure companies' sustainability in these regards. Nowadays, the inclusion of these concepts has become the norm rather than the exception in the corporate world.

Sustainable investments focusing on ESG is an exceedingly fast-growing market in the financial world. Morningstar (2021) reports that the amount of U.S exchange-traded funds focusing on sustainability increased fivefold in the past decade alone. According to GSIR (2020), global sustainable investments exceeded 35$ trillion in 2020, up 55% from four years prior, and in the U.S, the sustainable investments have doubled in the same period. A development that perhaps suggest a divergence away from a pure profit motive to a more "doing well while doing good" point of view. However, at the other end, one of the largest banking groups in the world, BNP Paribas, published a study reporting that global green financing, which aims to finance sustainable projects, has increased more than 100 times in the past decade where green borrowing went up from $5.2 million in 2012 to $540.6 million in 2021 (Reuters 2022). The substantial expansion this sector continues to have makes it a relevant topic to keep on researching.

Earlier studies on ESG investments and their implication for predicted future returns have seen varying results across studies using different methods, assets, and time periods. Kempf and Ostoff (2007) and Statman and Glushkov (2009) for example suggest that portfolios including

higher ESG rated firms have produced abnormal positive returns while according to Borgers et al. (2013) and Halbritter and Dorfleitner (2015) these findings may have been significant in the past but are non-significant now. Hong and Kacperczyk (2009) on the other hand claims that having an investment strategy that is more constrained on social norms could lead to negative abnormal returns.

By using one of the most extensive datasets to date on multiple markets, this essay attempts to shed more light on the literature's most frequently asked question: Does an investing approach in stocks based on ESG ratings lead to anomalous performance? The results given here are based on Refinitiv ESG rating[1]. The ratings are then used to construct portfolios of stocks with high (low) ratings as well as a difference portfolio combining the two. The performance of these portfolios is studied over the period 2006-2021, where the performance is measured using Carhart's (1997) four-factor model, an extended version of the Fama and French (1993) three-factor model, including a momentum factor. A series of modifications such as time, weighting, size, and selection criteria of the portfolios are added to control for robustness.

Previous literature focuses heavily on the U.S market, and one of the purposes of this paper is to build on former research. Adding additional markets in the form of the European market and emerging markets allows us to fill a knowledge gap and contribute to the literature. We are also using the most recent time period to replicate similar methods used in the literature to investigate if the conclusions are still valid today. As multiple researchers suggest, the results have changed over time, which may be of great interest now, following the rate at which the market for ESG investments has increased only in the past few years (Dorfleitner, Kreuzer & Sparrer, 2020; Pedersen, Fitzgibbons & Pomorski, 2021).

The remaining parts of the study have the following outline: Background and previous literature are reported in Chapter 2. Chapter 3 discusses the data collection and delimitations while Chapter 4 presents the methodology of constructing and evaluating the portfolios. The results of the study are presented and discussed in Chapter 5 and finally in Chapter 6 the conclusions are laid out.

[1] Previously known as ASSET4
https://www.esade.edu/itemsweb/biblioteca/bbdd/inbbdd/archivos/Thomson_Reuters_ESG_Scores.pdf [Last Accessed 10 August 2022]

2. Background and Previous Research

Chapter 2 gives a brief history of corporate responsibility and an overview of the past literature on ESG investing, namely through ESG portfolios. The sections start off by giving a background of the progress the role of sustainability has had in relation to corporate performance over the years, then continue to define the three pillars that ESG is built upon and what link these have had between firms' social and financial performance. Further, I go through the results of similar studies in past literature regarding the three regions selected and finish by giving the theoretical framework for the asset pricing model used in the study to evaluate the portfolios.

2.1 Corporate Social Responsibility and ESG

As previously mentioned, the connection between corporate responsibility and corporate performance is not something new. Some of its basic ideas in business ethics may even go back to Adam Smith, as argued by Whittaker (2011). The modern concept, however, comes in the form of the term CSR which Bowen (1953) introduces in his book. Here he discussed the topics of what reasonable social responsibility expectations one might assume on corporations, what steps might be taken to broaden the concept within corporations, and how the advances in this field could benefit society. Half a century later, in 2005, the United Nations started a process aimed at the world's largest institutional investors to try to integrate a set of Principles for Responsible Investments (PRI), leading to the integration of ESG into the decision-making when investing. At the time, the number of companies that had signed the principles amounted to 181 signatories with under $5 trillion in assets under management. In 2021, those numbers had turned to 3826 signatories and over $121 trillion in assets under management (PRI 2022). The increasing move to more sustainable investments by institutions led to an increased interest in the financial-driven aspect of ESG, increasing the attention in academia trying to link the relationship between firms' social and financial performance (Eccles, Lee & Stroehle, 2020)

2.2 The Pillars of ESG

In the same way that there are multiple interchangeable terms for the concept of firms' social performance, the way to measure it seems to differ just as much. Here an attempt is given to

describe the basic idea of the core measurement concepts for each individual pillar, as well as some previous research focused on specific pillars. At the beginning of each section, the pillars are briefly defined according to Refinitiv. More regarding Refinitiv's definition and rating system are discussed later in the data section.

2.2.1 Environmental

Refinitiv divides the environmental pillar into three categories: Emissions, including themes such as carbon emission and waste. Innovation involves green revenues, research and development (R&D), and capital expenditures, and finally, resource use in the form of water and energy.

The environmental pillar has seen growing interest as of late, something that Busch and Hoffmann (2011) attribute to publicly higher recognition and concern of the environmental change taking place in the world with issues such as climate change. In an attempt to measure the link between environmental issues and financial performance in corporations, Busch and Hoffman (2011) collected data regarding companies' carbon emissions from Sustainable Asset Management (SAM) which is one of the many organizations supplying ESG ratings. They extend their data by sending out a questionnaire to firms to measure corporations' internal efforts of an environmental-based strategy. From the regression results using their two different measurements, Busch and Hoffmann (2011) highlight the importance of identifying which kind of environmental data researchers use when comparing corporate environmental performance (CEP) and corporate financial performance (CFP) as they can be divided into two dimensions: Process-based CEP and Outcome-based CEP. Depending on what screening is used, the effects that should be considered and the outcomes to be anticipated alter. The process-based measurement indicates a negative relationship between environmental and financial performance, while the outcome-based using carbon emission data argues for a positive relationship.

Endrikat, Guenther and Hoppe (2014) build upon previous research in this pillar, including the study by Busch and Hoffmann (2011). Finding that the link between CEP and CFP is mainly positive but also that the causation here is bidirectional, meaning that there also is a positive relationship the other way around, from CFP to CEP. Pedersen, Fitzgibbons and Pomorski (2021) measure the environmental pillar or how green a company is on a proxy of carbon intensity based on emissions divided by sales. Creating portfolios based on this proxy, they find

no conclusive significant result for this ESG pillar. Kempf and Osthoff (2007) also create portfolios based on high and low-rated firms regarding the environment, finding some significance for positive abnormal returns in the high-rated portfolio.

2.2.2 Social

The social pillar is split into four different categories by Refinitiv (2022a): community, human rights, product responsibility, and workforce. The product responsibility category includes themes like responsible marketing, data privacy, and product quality. Workforce includes matters such as diversity and inclusion, working conditions, health and safety, as well as career development and training.

Kempf and Osthoff (2007) apply the same four categories seen in Refinitiv's social pillar, finding significant positive alphas in the high-rated portfolios for the human rights and community categories. When applying a best-in-class approach, taking the best firms in different industries, the significance level increase in the same two categories.

Hong and Kacperczyk (2009) take another approach to the social pillar, which involves studying "sin" stocks that include firms associated with alcohol, tobacco, and gambling products. They explain that two other industries are also commonly linked as immoral, the sex industry, and the defense industry. However, both are excluded from their selection of sin stocks as the first lack enough publicly traded firms and the defense industry because of the unclear view of this industry being classed as sinful in the U.S. They find that stocks included in their selection of sin stocks outperform other comparable stocks with a higher expected return. Concluding that being more constrained by social norms as an investor reflects negatively on returns. Pedersen, Fitzgibbons and Pomorski (2021) use the same concept of sin stocks in their measurement of the social pillar giving stocks related to the sinful industries a value of one and a value of zero if the firms are regarded as non-sin. Using difference portfolios weighted both equally and based on value, they find very little evidence for abnormal returns overall. Some proof of the results reported by Hong and Kacperczyk (2009) can be found using weighting based on market cap and applying some asset pricing models with few factor loadings.

2.2.3 Governance

The third and final pillar is the governance pillar consisting of three ESG categories being, management, shareholders and CSR strategy. Measuring management as compensation and structure (committees, independence, diversity), Shareholder by shareholder rights and takeover defenses. Last CSR strategy is based on ESG reporting and transparency as well as CSR strategy (Refinitiv 2022a).

By using the second category to connect corporate governance and firm performance, Gompers, Ishii and Metrick (2003) find that firms having more substantial shareholder rights have greater abnormal returns than firms with weak shareholder rights. Bhagat and Bolton (2008) question these results and, by using various measurement techniques, fail to find any correlation between corporate governance and future stock returns.

In forming ESG portfolios, Pedersen, Fitzgibbons and Pomorski (2021) measure the governance pillar by companies' accruals, defined as accounting income where the associated cash has not been obtained yet. Fewer accruals for a firm give a higher governance rating. He bases this measurement on the fact that low accruals suggest that the management of corporations accounts for their profits more conservatively, and a more cautious approach to the accounting process tends to be implemented by better-governed firms. By applying this method, they find highly significant positive alphas by taking a long position in firms with high governance ratings and a short position in low-rated firms. Other ESG portfolio studies such as Halbritter and Dorfleitner (2015) and Statman and Glushkov (2009) report nonsignificant abnormal returns for the governance pillar.

2.3 ESG Portfolios

Orsato et al. (2015) and Friede, Busch and Bassen (2015) illustrate a rapid development of the number of empirical studies on the topic of ESG performance during the last decades, which should be no surprise given the substantial growth in sustainable assets during the same period. The most prevalent approach in the literature to study the link between companies' social and financial performance has been by building ESG portfolios. Given that this is the approach used in this study, this chapter primarily focus on past literature that employed this method. ESG portfolios allow the application of different asset pricing models by combining extensive panel

data sets and reducing them into one-dimensional time series based on ESG ratings. The ESG portfolio approach typically involves creating one portfolio consisting of highly rated firms and one consisting of lowly rated firms and then adding a portfolio which is the difference between the former ones (Halbritter & Dorfleitner 2015). The results and presence of ESG investment studies vastly differ across regions (Friede, Busch & Bassen 2015).

2.3.1 ESG Portfolios U.S

Most of the past literature on the topic heavily concentrates on developed economies, with a large portion contributed to the U.S market and their firms (Cheng, Ioannou & Serafeim, 2014). While the size of the U.S market undoubtedly plays a role in this regard, another critical factor is due to the availability of data on ESG ratings. The ESG scores from Refinitiv, used here, only go back to 2002, and then it was mainly U.S and European coverage (Refinitiv 2022). One of the first data suppliers of ESG ratings were KLD Research & Analytics (KLD), providing numerical ESG scores for U.S firms going back to 1991 and has since been commonly used in the literature regarding ESG portfolios (Eccles, Lee & Stroehle, 2020; Halbritter & Dorfleitner, 2015)

Using KLD data for U.S firms, Kempf and Osthoff (2007) and Statman and Glushkov (2009) are two of the first studies that apply a similar methodology to this paper. They create and compare portfolios based on ESG ratings using the Carhart (1997) four-factor model in the sample period 1992-2004 and 1992-2007, respectively. Both use equally weighted portfolios, while Kempf and Osthoff (2007) also include a value-weighted approach based on the firms' market capitalization. However, Statman and Glushkov (2009) disclose that their method of only analyzing equally weighted portfolios is a deviation from practical portfolios for investors as these usually more resemble being value-weighted. They conclude in their studies that there are significant positive abnormal returns to be made by being a socially responsible investor and implementing an investment strategy based on ESG ratings, getting the best result from a best-in-class screening.

In more recent history, the literature suggests more insignificant results incorporating the ESG portfolio model in the U.S. Borgers et al. (2013) document that the trading strategies that generated positive anomalous risk-adjusted returns during 1992-2004 were essentially non-significant in the following period from 2004-2009. Pinpointing that these results largely fall in line with economic theory such as the efficient market hypothesis that mispricing in the market, which causes abnormal returns, is short-lived as the market absorbs and adjusts according to

information relevant to mispricing. Halbritter & Dorfleitner (2015) follows this up with a similar conclusion in perhaps one of the more extensive research projects on the ESG portfolio method in the U.S as they add two more sources of ESG ratings in their study. Their study's findings heavily question the suggested results in previous literature of positive abnormal returns using an ESG investment strategy.

2.3.2 ESG Portfolios Europe

In the literature, Europe seems to be the second most studied region for ESG investments (Friede Busch & Bassen, 2015). However, in terms of the size of the global ESG market, Bloomberg (2021) states that Europe remains by far the biggest and most developed region in this specific market and that in the third quarter of 2021, Europe accounted for more than three-quarters of the global inflows of investments directed towards ESG. Previous literature on ESG-based investments reports a less appealing result for the practical investor.

In their article Auer and Schuhmacher (2016) report that ESG-based investment strategies not only find a lack of significant evidence on their outperformance of standard investment strategies but that investing responsible comes with a price ending up with negative abnormal risk-adjusted returns in comparison to the benchmark. Bannier, Bofinger and Rock (2019) found similar results by studying ESG portfolios both in the U.S. and Europe between 2003-2017, suggesting that in both these markets, lower ESG-rated firms yield strong positive returns. They continue that these positive returns come with a risk premium, i.e., that portfolios with high ESG scores have reduced risk but that this risk insurance is more attributed to U.S firms rather than European firms.

These results are further affirmed by a recent study, as Dorfleitner, Kreuzer and Sparrer (2020) indicate that significant positive abnormal returns were found in Europe by investing in a portfolio of the worst rated ESG firms and a high-low portfolio resulted in significant negative returns in the subperiod between 2002-2010, both results turned non-significant in the latter period between 2010-2018 indicating that the negative abnormal returns might be a result of the past.

2.3.3 ESG Portfolios Emerging Markets

The basic concept of the term emerging markets is that these economies or countries belong to a group in a transition phase of an emersion from being a developing country to becoming a part of the group of developed economies (Bekaert & Harvey, 2002). The countries included in the emerging markets group differ between institutions. This study uses stocks from the 24 countries listed in MSCI emerging market index (MSCI 2022).

As suggested by Cheng, Ioannou and Serafeim (2014), previous academia on the link between corporate responsibility and corporate performance has heavily concentrated on developed economies like the U.S and Europe. Emerging markets, which represent a large quantity of business globally, appear to have been left out in the fast-growing academia that revolves around ESG investment (Orsato et al. 2015; Garcia, Medes-Da-Silva & Orsato, 2017).

Studies incorporating the same methodology in the form of ESG portfolios seem to be even scarcer. Some have investigated the effects of ESG strategies in emerging market indices (Pollard, Sherwood & Klobus, 2017). They note that based on the historical progress of ESG integration and as the literature continues developing and more data becomes available, ESG investments may potentially bring higher risk-adjusted returns in emerging markets. In the study by Friede, Busch and Bassen (2015), which summarizes a large quantity of ESG-related research, they suggest, without going into much detail, that alongside North America, emerging markets are a sub-region where opportunities for ESG abnormal returns exist.

The evident lack of comparable studies for this market motivates an opportunity to fill a research gap in the literature. Historically, as developing and emerging markets may have had fewer social and environmental needs and interests than developed markets (Dobers and Halme 2009). It may be of more importance to study ESG in these markets. A simple justification for the lack of studies can be explained by the absence of ESG data in these countries. As Refinitiv now provides more reliable data in this market for a reasonable sample period, this offers an appropriate timing for implementing ESG portfolios in emerging markets.

2.4 Asset Pricing Models.

The modern portfolio theory established by Markowitz (1952) had its core concept in investors being risk averse. If given the choice of portfolios, the investor would minimize the variance or risk of portfolio return given expected returns and maximize expected return, given the variance. As a result, his method is usually referred to as a mean-variance model. This portfolio selection approach by Markowitz (1952) set the foundations for asset pricing theory by Shape (1964) and Litner (1965) with the construction of the Capital Asset Pricing Model (CAPM), which has been consistently used in academia ever since. CAPM extended the ideas of Markowitz (1959) and is formulated in the following equation:

$$r_{i,t} - r_{f,t} = \beta_i(r_{m,t} - r_{f,t}) \quad (1)$$

Where $r_{i,t} - r_{f,t}$ is the expected return of the asset over the risk-free rate. $(r_{m,t} - r_{f,t})$ denotes the market factor and β_i is the investments beta.

Although praising its fundamental concept and building ground for more advanced models, Fama and French (2004) warn about the model's empirical problems limiting its applicability. Suggesting that a univariate model is not enough to explain average stock returns by the estimated beta accurately, something that inserting additional variables and factors had suggested in research in the late 70s. Concerning their criticism of the univariate model, Fama and French (1993) expand on the CAPM. Here they present two new explanatory variables of expected stock returns known as the SMB and the HML factors. The SMB factor stands for small minus big and relates to company size. This factor takes into account that small-cap companies have historically performed better than large-cap companies in terms of excess returns (average returns). The second introduced factor, HML, which stands for high minus low, refers to a company's book-to-market ratio. This factor is based on the observation that high book-to-market companies, also called "value stocks," typically generate higher returns than "growth stocks" that represent low book-to-market companies.

Combining the two above-mentioned risk factors with the market factor, this model is known as the Fama and French (1993) three-factor model or for short FF3, formulated as:

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_i(r_{m,t} - r_{f,t}) + \beta_i SMB_{i,t} + \beta_i HML_{i,t} + \varepsilon_{i,t} \quad (2)$$

Where α_i describes Jensen's alpha measuring abnormal performance, the size and value factor is denoted by $SMB_{i,t}$ and $HML_{i,t}$ respectively. The betas or β_i are the estimated coefficients for the different factors and $\varepsilon_{i,t}$ is the error term or the residuals. Everything else is the same as the CAPM model.

Based on a momentum abnormality noted by Jegadeesh and Titman (1993), Carhart (1997) expanded the model further by developing a four-factor model which incorporates a momentum factor called PR1YR, into the Fama and French (1993) model. He defined and calculated this momentum factor as the difference between the firms with the top 30% eleven-month returns, which lagged one month, minus the bottom 30% performing firms of eleven-month returns lagged one month. This four-factor model was found to reduce the average pricing errors in the CAPM and FF3 model, demonstrating that it better captured average stock returns on a cross-sectional basis

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_i\left(r_{m,t} - r_{f,t}\right) + \beta_i SMB_{i,t} + \beta_i HML_{i,t} + \beta_i PR1YR_{i,t} + \varepsilon_{i,t} \qquad (3)$$

Where $\mathrm{PR1YR}_{i,t}$ is the momentum factor, everything else is the same as the FF3 model.

As previously discussed, the four-factor regression by Carhart (1997) has been the model of choice in similar research (see Halbritter and Dorfleitner, 2015; Kempf and Osthoff, 2007; Statman and Glushkov, 2009). To better relate and compare the results of this study with earlier literature, this model is also selected as the primary model.

3. Data and delimitations

This section sets out to report the data used in this study, giving an explanation of what data is collected to use for the results illustrating some descriptive statistics on the data in figures and tables. The second part of the chapter mentions different limitations this study has, based on the collected and not collected data.

3.1 Ratings Data

ESG scores used for this study were collected from Refinitiv through Thomas Reuters Eikon. Refinitiv collects publicly available ESG information based on reported data from companies such as annual reports, news sources, and stock exchange filings. This information is then aggregated to measure organizations' ESG performance across ten ESG category scores which are then compounded to build the three pillars E, S, and G (Refinitiv 2022a). The categories for the respective pillars are illustrated in figure 1. This results in each pillar getting a score between 0 and 100.

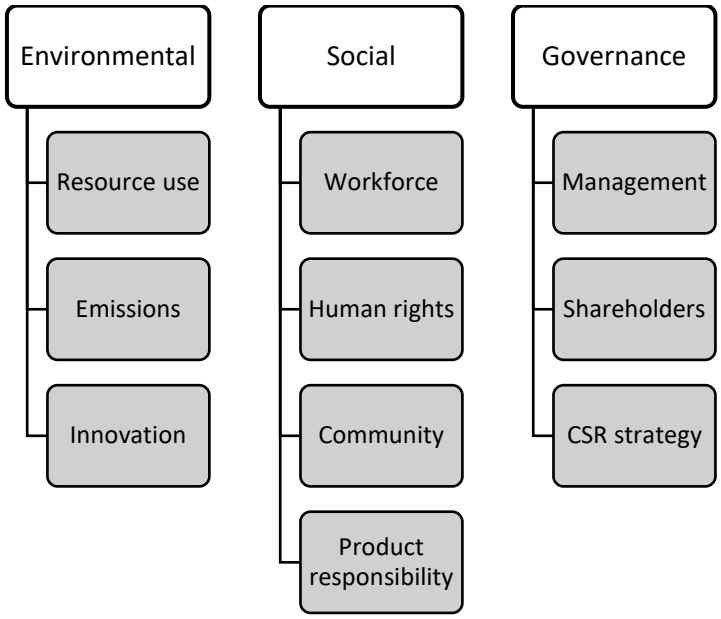

Figure 1: This figure shows the ten ESG categories attributed to the three pillars, which are used to measure the score of the individual pillars

The next step is to calculate the overall ESG score which is done by a weighted sum of the three pillar scores (Refinitiv 2022a).

When collecting the ESG data for each year in the sample period, a new criterion was implemented each year, meaning that if there was ESG data for a company for one year, that company is included. This means that the sample data includes firms that did not have ESG data at the beginning of the sample but also includes ESG data that did have ratings earlier in the period but not in the present due to removal, insolvency, or merger. This way, the paper eliminates being subject to a common bias called survivability bias. Table 1 illustrates the descriptive statistics of the ESG rating sample across the markets. In most of the categories for the three observed markets, a small increase in the mean and standard deviation can be seen over time from the starting year of the sample period, indicating that, on average, corporations may have become more socially responsible. The increase in standard deviation can also be an effect of the increasing number of firms covered by ESG rating agencies over time. A noteworthy result is that for many categories, the mean value seemed to increase from the middle of the period in 2013 to the end of 2020. Two possible explanations for this are the large increase in the number of firms included but also may be due to regulations for what constitutes a higher rating in ESG getting stricter through the years

As shown in the table, the environmental pillar is the only one with a minimum value of zero. This is not because Refinitiv's system has assigned a company a zero rating directly based on company information but more due to the lack of information given by the company on this pillar. Refintiv's rating system is based on automatic data processing to avoid subjectivity. However, this also causes a company that has not reported information on the criteria for the environmental pillar to receive a value of zero, defined as a missing value. While this could correlate to their actual value if they had reported, it still makes the reliance on the ESG rating for these companies more questionable, as suggested by Sahin et al. (2022). Due to this, an additional portfolio is added for the E pillar. This means there are two low environmental portfolios, one where the zero values are included and one where they are excluded. This is only done for the original approach in the full period for the three markets. Additional portfolio constructions for the environmental pillar exclude the firms that have zero values as these values may not accurately correlate with these firms' environmental efforts (Sahin et al. 2022)

Table 1

Descriptive statistics of the ESG ratings

		Year	Mean	Min	Max	Sd	Obs
U.S	**ESG**	2005	34,01	3,71	86,29	16,78	461
		2013	43,60	0,92	92,34	19,75	805
		2020	39,24	1,05	93,62	18,80	3416
	ENV	2005	18,37	0,00	90,54	22,82	461
		2013	35,16	0,00	98,51	27,99	805
		2020	22,20	0,00	97,25	26,72	3416
	SOC	2005	35,29	2,81	97,20	19,33	461
		2013	45,68	1,19	95,73	21,21	805
		2020	42,77	0,31	97,68	20,70	3416
	GOV	2005	47,59	3,21	96,30	21,45	461
		2013	49,13	0,45	96,70	23,08	805
		2020	46,69	0,65	99,56	22,35	3416
Europe	**ESG**	2005	37,87	5,36	89,16	18,27	572
		2013	51,23	0,63	94,49	20,32	824
		2020	49,08	1,57	95,22	22,06	2357
	ENV	2005	29,96	0,00	97,65	26,52	571
		2013	52,41	0,00	98,58	26,88	824
		2020	41,52	0,00	99,20	28,15	2357
	SOC	2005	38,27	3,17	98,03	19,95	571
		2013	51,69	0,12	97,99	24,00	824
		2020	51,33	0,44	97,14	24,72	2357
	GOV	2005	46,90	1,12	99,00	22,86	572
		2013	50,30	1,38	97,44	22,27	824
		2020	51,49	0,94	97,27	23,31	2357
Emerging Markets	**ESG**	2005	32,53	14,19	60,85	14,00	27
		2013	40,19	1,56	91,12	21,46	883
		2020	45,20	1,57	93,55	20,89	2082
	ENV	2005	26,62	0,00	60,41	19,72	27
		2013	35,79	0,00	95,61	26,53	883
		2020	38,87	0,00	97,84	26,88	2082
	SOC	2005	30,66	4,69	77,11	19,51	27
		2013	38,37	0,15	97,15	25,02	883
		2020	44,96	0,50	97,70	26,02	2082
	GOV	2005	44,97	16,38	78,89	16,49	27
		2013	48,65	1,03	98,02	23,17	883
		2020	49,89	0,54	95,36	22,01	2082

This table presents some descriptive statistics on the ESG ratings data, reporting values like the mean, minimum, maximum, standard deviation and the number of observation for 3 different years over the sample period.

The reasoning for examining the three selected markets mainly comes from both wanting to include a market like the U.S to compare to earlier research but also a region like emerging markets where there is a clear lack of previous literature. However, the selection partly stems from trying to find financial regions with somewhat equal sizing and where the development of the number of rated ESG firms by the agency was approximately equal over time across the regions. This development is demonstrated in figure 1, where the number of firms rated by Refinitiv increased from around 1000 at the start of the sample period to around 8000 at the end.

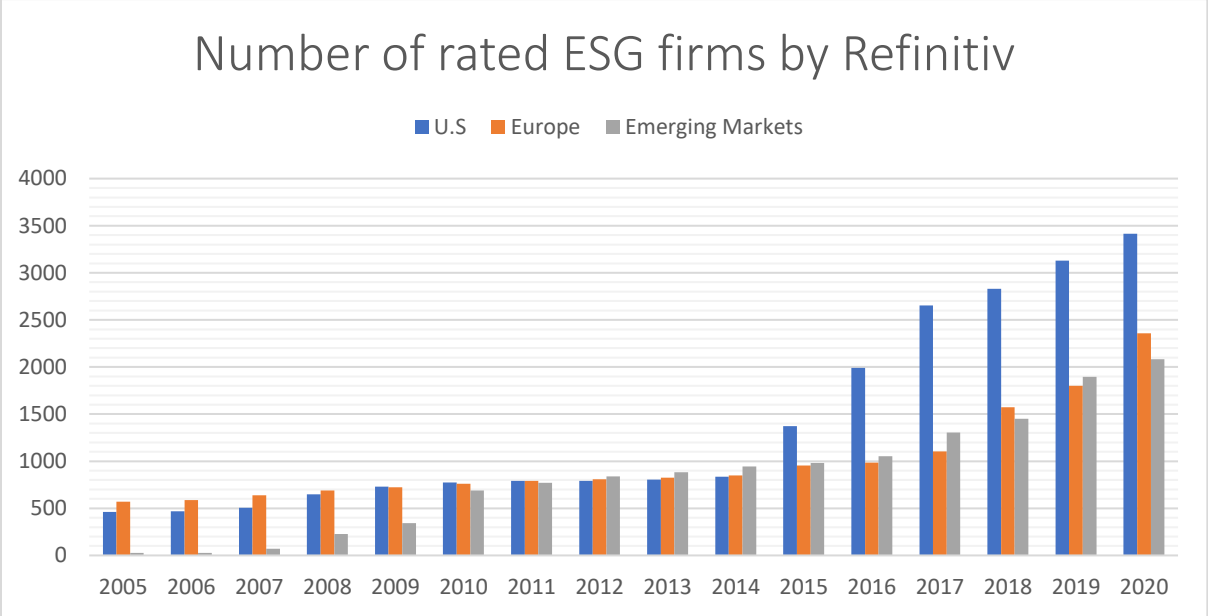

Figure 2 demonstrates the development in the number of rated firms by Refinitiv for the three selected markets over the sample period

Market capitalization was also gathered using Refinitiv for every firm each year between 2005-2020 like the ESG ratings, while the monthly return data was collected through FactSet retrieving stock closing price each month between 2006-2021.

The logic for the difference in the sampling period is that at the beginning of each year, the ESG portfolios are selected based on the ESG ratings at the year's end. Meaning that the ESG ratings at the end of 2005 were used to select companies in the portfolios for the coming year, meaning that January 2006 is the first month returns are needed. The reason for this specific period comes mostly down due to the availability of Refintiv's database, which is undoubtedly the case for the emerging markets, but also because the findings are supposed to be recent and up to date to better represent the current conditions of the stock market.

3.2 Delimitations

- Due to the lack of ESG-rated firms in emerging markets at the beginning of the sample period, there is a delay in the start period for this region in the study. It is instead starting from 2009. This means that this market is disregarded in the first subperiod in the robustness regressions.

- There are multiple suppliers of ESG ratings, as mentioned in the chapter on previous literature. While the standard method in past studies has been to use only one source, some researchers express the potential dangers of this due to the low correlation in ratings between agencies (Berg, Kölbel & Rigobon, 2019; Halbritter & Dorfleitner, 2015). The reason for only using one provider is accessibility and time limitations.

- In prior research of ESG portfolios, several different asset pricing models have been used, including CAPM, the Fama and French (1993) three-factor model, and the Fama and French (2015) five-factor model. The Carhart (1997) four-factor model is the only regression model applied here but is also one of the most commonly used, especially regarding ESG portfolios referenced in this study, which this study mainly compares with.

- Possible bias due to the steep increase in the number of stocks included over the sample period. Kempf and Osthoff (2007) comment on this potential bias as one of the arguments to only include stocks covered in the KLD database over the whole sample period. They also report that they control for such bias by re-estimating the results using all stocks, finding no prominent difference. However, as this paper aims to reflect a simple and practical portfolio construction method any investor could follow over time, all stocks with an ESG rating at the end of the calendar year are included in the following year's rebalancing.

- Transaction cost. Since there is a rebalancing of the stocks in the portfolios each year, one must consider the turnover ratio and the transaction cost that follows with buying and selling stocks. Transaction costs are discarded in this paper but should be considered when interpreting the results.

- Missing values, for a few companies the return data was not available at the start they had ESG ratings meaning that for brief periods some portfolios may have one or two less active holdings, while it should have no direct impact on the portfolios particularly the larger one it could in some instances affect the smaller portfolio size. If there is missing values in a portfolio, the weighting gets transferred to the rest of the firms.

4. Method

Chapter 4 lays out the process in constructing the ESG portfolios and presenting the different types of portfolios that are created in order to perform robustness checks. This is followed by presenting the selected model that is used to evaluate the portfolios.

4.1 Portfolio Construction

The first step is to calculate the companies monthly return using their monthly closing price data computed by the standard equation

$$R_{i,t} = \frac{P_{i,t} - P_{i,t-1}}{P_{i,t-1}} \quad (4)$$

$R_{i,t}$ denotes the return of the stock for firm i at time t while $P_{i,t}$ represent the price of stock for firm i at time t and $P_{i,t-1}$ the price for the month prior to time t.

The building of the portfolios closely complies with the theoretical framework for ESG portfolios described in previous chapters. The first step taken is to construct the High/Low portfolios, which is done by taking the 100 best/worst performing firms in a particular category. The portfolio size is somewhat arbitrary but chosen to some extent to resemble the number of holdings in a fund. Jiang, Yao and Yu (2007) aggregated the number of holdings held across more than 2000 U.S funds between 1980-2002 to be 102,15 stocks. However, to investigate the dependency of the portfolio size on the results, a portfolio of the 20 best/worst performing firms are also added. A size that might more represent the average investor and needed for a diversified portfolio (Chong & Phillips, 2013)

4.1.1 Difference Portfolios/High Minus Low Portfolios

As the objective is not only to study how the high (low) portfolios perform with the market but also how they perform in relation to each other, difference portfolios or high minus low portfolios are created. In theory, this is done by taking a long position in the portfolio with high-rated firms and an equal short position in the low-rated portfolio. This paper creates this portfolio by subtracting the low-rated portfolio from the high-rated portfolio and then applying the Carhart (1997) four-factor model to check if there are any abnormal returns.

$$r_{H,t} - r_{L,t} = R_{HML,t} \quad (5)$$

Where $r_{H,t}$ represents the return of the portfolio with high ESG ratings and $r_{L,t}$ the return of the portfolio with low ratings. $R_{HML,t}$ is then the return of the difference or high minus low portfolio

4.1.2 Robustness Checks

The initial portfolios use an equal weight for all the holdings in the portfolio which in this case means 1% each and constructed using the following formula

$$R_{P,t} = Average\left(\sum_{n=1}^{n} R_{i,t}\right) = \frac{\sum_{n=i}^{n}(R_{i,t})}{n} \quad (6)$$

Where $R_{P,t}$ is the return of the portfolio, $R_{i,t}$ is the return of stock i and n equal the number of stocks included in the portfolio.

To control the robustness of the results, a series of additional portfolios are constructed. One of them is to use the market capitalization of the firms to weigh the size of the holdings in each firm. Where you sum the market capitalization for all firms in the portfolio and then divide each individual firm with the sum to get the individual weights.

Every year between 2006 to 2021, two portfolios are constructed for each ESG category, meaning that there is a rebalancing of the portfolios at the start of every year. The rebalancing is made based on the companies' ESG rating at the end of the year before. This process is the same for the portfolios that are not equally weighted. They are then using the market capitalization of the companies at the end of the year before when constructing the weighted portfolios.

As previously discussed and illustrated, a delimitation of the study is how far back the historical ESG rating data goes. For the emerging markets, Refinitiv does not have enough rated firms to construct the high/low portfolio until 2009 (using 2008 ESG data), which is why these portfolios have a later start period. In examining the connection between the social and financial performance over time, the data set is, in addition to the full period, split into two equally long subperiods, one from 2006-2013 and the other between 2014-2021.

4.1.3 Best-In-Class Approach

This paper also includes a portfolio creation based on a best-in-class approach, meaning that companies are chosen for the portfolios based on their ESG performance relative to a specific economics sector. Using The Refinitiv Business Classification (TRBC), companies are classified based on a hierarchal structure containing five levels sorting the firms into 13 economic sectors, 32 business sectors, 61 industry groups, 153 industries, and 895 activities (Refinitiv 2022b). Three of the 13 economic sectors were added in recent years, including institutions, association and organizations, government activities, and academic & educational services. Of the three recently added to the TRBC system, only academic & educational services are included in the sample data. However, due to the limited number of firms in this sector included in the sample. These firms were excluded from this approach. Instead, this study considers ten economic sectors in the portfolio selection. The different sectors are basic materials, industrials, consumer cyclicals, consumer non-cyclicals, financials, healthcare, technology, utilities, and real estate. Both Kempf and Osthoff (2007) and Halbritter and Dorfleitner (2015) include ten different industry classes in their best-in-class approach. However, there are some slight differences in the classifications of the sectors between the studies.

This paper includes the top (bottom) performing firms in each economic sector based on a fixed number of firms amounting to the portfolio sizes of 100 and 20 firms used in previous portfolio applications. This means that the top (bottom) 10 (2) firms in each specific sector are selected in the portfolios and rebalanced at the start of every year. In other aspects, these portfolios are measured in the same way as previous ones, with two different portfolio sizes, both equally- and value-weighted, estimating both the whole period and the subperiods.

The argument for including this approach is to strengthen the results by avoiding a possible bias in the selection where for certain pillars or overall, the high (low) rated firms in the ESG portfolios are over-represented by a particular economic sector (Kempf & Osthoff, 2007)

4.2 Portfolio Evaluation

As previously discussed, the four-factor model by Carhart (1997) has been a frequent model of choice in similar research. To be able to better relate and compare the results of this study with earlier literature the four-factor model is selected as the primary model and is estimated using the following equation:

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_i(r_{m,t} - r_{f,t}) + \beta_i SMB_{i,t} + \beta_i HML_{i,t} + \beta_i MOM_{i,t} + \varepsilon_{i,t} \quad (7)$$

Where $r_{i,t}$ is the expected return of the investment and $r_{f,t}$ is the risk-free investment option. α_i is Jensen's Alpha measuring abnormal performance which is performance not clarified by the factor loadings. β_i denotes the estimated coefficients for the different factors. The different factors, market, size, value and momentum are in their turns represented by $(r_{m,t} - r_{f,t})$, $SMB_{i,t}, HML_{i,t}$ and $\mathrm{MOM}_{i,t}$ respectively. Regarding the momentum factor, this was in the literature review referred to as PR1YR but for simplicity, it is mentioned as MOM from now on as this notion is on the table results. Finally, $\varepsilon_{i,t}$ is the error term or the residual.

The risk factors needed for the Carhart (1997) four-factor model were taken from Kenneth R. French data library[2]. As this essay is using three different markets, three sets of factor loadings are needed to comply with each market.

Excel was used to handle all the collected data, as well as to perform the portfolio calculations, including the multivariate OLS regression from the four-factor model.

[2] https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html (Last accessed on 2022-08-15)

5. Results and Discussion

Chapter 5 demonstrates the results estimated by using the data collected and applying the chosen methodology, reporting the results in several tables of summarized regressions. This chapter is divided into five parts, one for each selected market, one including all the robustness checks implemented, and a final one that gives a general discussion of the results and some suggestions for further research.

5.1 U.S. Market Results

Table 2 demonstrates the result of the time-series regression based on 14 equally weighted and 14 market cap weighted ESG portfolio strategies for the U.S market.

The alphas are analyzed to estimate any potential abnormal performance for the portfolios. If the value on the alpha is positive, the portfolio is outperforming the market benchmark, while a negative alpha indicates that the strategy is underperforming. When comparing the two high and low portfolios in a category, the sign of the difference portfolio, high minus low, suggests which portfolio performs best in each category. i.e., a positive score on the high minus low portfolio means that the high portfolio performs better than the low portfolio and vice versa. P-values of the alphas then indicate whether these possible abnormal performances are statistically significant.

At first glance, looking at table 2, one can see that the sign of the alpha values on the difference portfolios are largely inversed between the equally weighted strategies and the portfolios that are weighted based on market cap. The portfolio containing firms with high environmental and social scores reports negative abnormal returns that are highly significant for both weighing methods. There are also strongly significant negative anomalous returns in the overall ESG portfolio using market cap weighting. In general, for all categories, the portfolios containing high ESG-rated firms show negative risk-adjusted returns where many values are significant. For the portfolios constructed on low-rated firms, a more positive outcome is estimated for those weighted equally.

Table 2 Equally and Market Cap weighted ESG portfolios for the U.S. market

Equally weighted							
		Alpha	MKT	SMB	HML	MOM	R^2
ESG	high	-0,001	1,012***	0,000	0,152***	-0,130***	0,963
	low	0,002	1,061***	0,813***	-0,066	-0,181***	0,807
	high-low	-0,003	-0,049	-0,813***	0,218***	0,051	0,318
ENV	high	-0,002***	0,995***	0,049	0,212***	-0,133***	0,961
	low^	0,001	0,987***	0,905***	-0,109	-0,151***	0,815
	high-low^	-0,003	0,008	-0,856***	0,321***	0,018	0,373
	low	-0,001	1,020***	0,547***	0,178***	-0,207***	0,925
	high-low	-0,001	-0,025	-0,498***	0,034	0,074**	0,414
SOC	high	-0,002**	1,000***	0,063**	0,043	-0,148***	0,963
	low	0,003	1,021***	0,721***	0,015	-0,145***	0,839
	high-low	-0,005**	-0,020	-0,658***	0,028	-0,003	0,264
GOV	high	0,001	1,028***	0,151***	0,252***	-0,157***	0,948
	low	0,001	1,119***	0,744***	-0,113	-0,167***	0,820
	high-low	0,000	-0,091	-0,593***	0,365***	0,010	0,247
Market Cap Weighted							
		Alpha	MKT	SMB	HML	MOM	R^2
ESG	high	-0,002***	0,945***	-0,253***	0,082***	-0,009	0,961
	low	-0,004**	0,935***	0,437***	-0,106	-0,129**	0,758
	high-low	0,002	0,009	-0,689***	0,188**	0,120**	0,263
ENV	high	-0,003***	0,947***	-0,210***	0,177***	-0,037**	0,961
	low^	-0,004	0,974***	0,331*	-0,162	-0,184*	0,411
	high-low^	0,001	-0,027	-0,540***	0,339**	0,148	0,050
	low	-0,008***	1,133***	0,212**	-0,175**	-0,180***	0,755
	high-low	0,005*	-0,186***	-0,421***	0,352***	0,144**	0,201
SOC	high	-0,002***	0,939***	-0,245***	0,019	-0,008	0,958
	low	-0,003*	0,918***	0,439***	0,013	-0,045	0,792
	high-low	0,001	0,021	-0,683***	0,006	0,036	0,284
GOV	high	0,000	0,991***	-0,293***	0,002	-0,006	0,915
	low	-0,002	1,051***	0,291***	-0,243***	-0,122***	0,857
	high-low	0,003	-0,060	-0,584***	0,245***	0,116**	0,279

*This table presents the results of the Carhart (1997) four-factor model over a sample period from 2006 to 2021 on a monthly basis. Regressions are run for each portfolio individually. The high (low) portfolios include the 100 companies with the highest (lowest) rating overall and for each individual category. The high-low portfolio represents the difference portfolios where a long position is taken in the high rated portfolio and a short position in the low rated portfolio. Alphas, factor loadings such as market (MKT), size (SMB), value (HML), momentum (MOM) and adj. R-squared are stated in the table. *.**.*** imply a significance level of 10,5 and 1% respectively based on the variables p-value. ^ means that firms having a value of zero in the environmental pillar is included*

5.2 European Market Results

Table 3 reports the same number of ESG portfolios for the European market. Similar to the results of the U.S. market, the market cap weighted version of the overall ESG portfolio with high-rated firms notes significant negative results at the 1% level. Other notions include that the low-rated portfolio consistently outperforms the high-rated portfolio using equal weights on

the firms. Here the low-rated overall ESG portfolio has positive abnormal returns significant at a 5% level. This significance level is also accurate for the negative abnormal returns in the difference portfolio in the same category.

Table 3 Equally and Market Cap weighted ESG portfolios for the European market

Equally Weighted							
		Alpha	MKT	SMB	HML	MOM	R^2
ESG	high	-0,001	1,044***	-0,068**	0,250***	-0,138***	0,980
	low	0,004**	1,082***	0,672***	-0,185*	-0,228***	0,846
	high-low	-0,005**	-0,038	-0,739***	0,436***	0,089	0,265
ENV	high	0,000	1,095***	-0,003	0,356***	-0,179***	0,981
	low^	0,001	1,110***	0,732***	-0,379***	-0,228***	0,856
	high-low^	-0,001	-0,015	-0,735***	0,735***	0,049	0,412
	low	0,001	1,109***	0,831***	-0,037	-0,182***	0,953
	high-low	-0,001	-0,014	-0,835***	0,393***	0,003	0,574
SOC	high	0,000	1,030***	-0,068*	0,196***	-0,128***	0,975
	low	0,004	1,107***	0,681***	-0,188*	-0,216***	0,853
	high-low	-0,004*	-0,077	-0,750***	0,385***	0,088	0,259
GOV	high	-0,001	1,044***	-0,068**	0,250***	-0,138***	0,980
	low	0,001	1,072***	0,731***	0,003	-0,151***	0,953
	high-low	-0,002	-0,028	-0,799***	0,248***	0,013	0,524
Market Cap Weighted							
		Alpha	MKT	SMB	HML	MOM	R^2
ESG	high	-0,003***	0,976***	-0,332***	0,127***	-0,020	0,980
	low	-0,001	1,116***	0,479***	-0,141**	-0,189***	0,927
	high-low	-0,002	-0,140***	-0,811***	0,268***	0,170***	0,464
ENV	high	-0,002***	1,000***	-0,271***	0,267***	-0,058***	0,982
	low^	-0,003**	1,111***	0,561***	-0,221***	-0,069	0,916
	high-low^	0,001	-0,111***	-0,832***	0,487***	0,011	0,515
	low	-0,002	1,115***	0,430***	-0,253***	-0,141***	0,932
	high-low	-0,001	-0,115***	-0,700***	0,520***	0,084*	0,486
SOC	high	-0,003***	0,976***	-0,353***	-0,015	0,006	0,978
	low	-0,004**	1,126***	0,494***	0,042	-0,058	0,890
	high-low	0,001	-0,151***	-0,847***	-0,057	0,064	0,385
GOV	high	0,000	1,044***	-0,185***	0,212***	-0,213***	0,966
	low	-0,003**	1,139***	0,488***	-0,122	-0,076	0,899
	high-low	0,003	-0,095**	-0,673***	0,335***	-0,137**	0,311

*This table shows the results of the Carhart (1997) four-factor model over a sample period from 2006 to 2021 on a monthly basis. Regressions are run for each portfolio individually. The high (low) portfolios include the 100 companies with the highest (lowest) rating overall and for each individual category. The high-low portfolio represents the difference portfolios where a long position is taken in the high rated portfolio and a short position in the low rated portfolio. Alphas, factor loadings such as market (MKT), size (SMB), value (HML), momentum (MOM) and adj. R-squared are stated in the table. *.**.*** imply a significance level of 10,5 and 1% respectively based on the variables p-value . ^ means that firms having a value of zero in the environmental pillar is included*

In general, the results of the ESG investments in the European market reaffirm to some degree the claims in past literature on Europe regarding high portfolios returning significant negative abnormal returns while the low in equally weighted returns positive alphas.

A common occurrence in both the two developed markets observed is that the alpha of the low-rated portfolio regularly outperforms both the high-rated portfolio but also the market benchmark in the equally weighted portfolios. This is, however, not significant in most categories. For the market cap weighted, the relationship is more inversed where the high-rated firms perform better than the low-rated, although both seem to perform worse than the benchmark as they frequently have a negative alpha.

5.3 Emerging Markets Results

Table 4 presents the results of the third and final market studied for the somewhat new market addition to the literature on ESG investing. While the estimations for emerging markets do not stand out much in terms of significant results. The portfolio strategies in this market seem to follow the theme of previous markets as the alphas for the low-rated portfolios are positive and higher than their high-rated counterparts when using equal weighting. The resemblance in result to the developed regions is repeated when looking at the portfolios weighted based on market cap. Decreasing the alphas of the low portfolios to slightly negative values, while the alphas are consistently just below zero for most of the high rated portfolios using both weighing methods. The finding for ESG portfolios in emerging markets here contradicts the suggestions made by fried et

Throughout the regression results in the three studied markets, the adjusted R-square value has been consistently high, suggesting that the Carhart (1997) four-factor model yields high accuracy in explaining the variation of the dependent variable being the portfolio returns.

Table 4 Equally and Market Cap weighted ESG portfolios for Emerging Markets

Equally Weighted							
		Alpha	MKT	SMB	HML	MOM	R^2
ESG	high	-0,001	1,002***	-0,069	0,433***	-0,173***	0,939
	low	0,003	0,953***	0,342***	0,043	-0,108	0,853
	high-low	-0,004	0,049	-0,411***	0,390***	-0,065***	0,157
ENV	high	-0,002	1,026***	-0,081	0,515***	-0,130***	0,934
	low^	0,002	0,973***	0,235**	-0,064	-0,070	0,869
	high-low^	-0,003	0,054	-0,316**	0,579***	-0,060***	0,266
	low	0,001	0,957***	0,304***	0,039	-0,058	0,889
	high-low	-0,003	0,069*	-0,385***	0,476***	-0,072***	0,278
SOC	high	-0,001	1,005***	-0,102	0,418***	-0,143***	0,935
	low	0,003	0,937***	0,274**	0,014	-0,097	0,828
	high-low	-0,003	0,068	-0,376**	0,404***	-0,045***	0,125
GOV	high	-0,002*	0,971***	0,049	0,331***	-0,099***	0,943
	low	0,001	0,995***	0,217**	0,091	-0,089	0,905
	high-low	-0,004*	-0,024	-0,168	0,240**	-0,010**	0,033
Market Cap Weighted							
		Alpha	MKT	SMB	HML	MOM	R^2
ESG	high	-0,002	0,898***	-0,355***	0,415***	-0,011***	0,887
	low	-0,001	0,880***	0,277**	-0,093	0,029	0,805
	high-low	-0,001	0,018	-0,632***	0,508***	-0,040***	0,202
ENV	high	-0,002	0,854***	-0,299***	0,533***	-0,009***	0,863
	low^	-0,001	0,934***	0,146	-0,272***	0,101***	0,829
	high-low^	-0,001	-0,080	-0,446***	0,805***	-0,110***	0,300
	low	-0,001	0,878***	0,214*	-0,125	0,048	0,809
	high-low	-0,001	-0,024	-0,514***	0,658***	-0,057***	0,202
SOC	high	-0,001	0,933***	-0,358***	0,399***	0,031***	0,897
	low	-0,001	0,881***	0,222	-0,065	0,069	0,740
	high-low	0,000	0,052	-0,579***	0,464***	-0,038***	0,143
GOV	high	-0,002	0,819***	-0,160	0,285***	0,016***	0,867
	low	-0,004**	0,979***	-0,040	0,047	0,021	0,869
	high-low	0,002	-0,160***	-0,120	0,238**	-0,004**	0,108

*This table shows the results of the Carhart (1997) four-factor model over a sample period from 2009 to 2021 on a monthly basis. Regressions are run for each portfolio individually. The high (low) portfolios include the 100 companies with the highest (lowest) rating overall and for each individual category. The high-low portfolio represents the difference portfolios where a long position is taken in the high rated portfolio and a short position in the low rated portfolio. Alphas, factor loadings such as market (MKT), size (SMB), value (HML), momentum (MOM) and adj. R-squared are stated in the table. *.**.*** imply a significance level of 10,5 and 1% respectively based on the variables p-value. ^ means that firms having a value of zero in the environmental pillar is included*

5.4 Robustness Testing

The previous section in this chapter looked at the results of the full period estimation using a straightforward method of selecting the top (bottom) 100 rated firms in each market. Here the results testing these estimations for robustness is presented, controlling the portfolios by size, period, and firm selection criteria in the form of the best-in-class approach.

The tables in this section only present the alphas estimating abnormal returns of the difference portfolios. Tables including the high (low) portfolios for the same robustness controls are found in the appendix.

5.4.1 Controlling for Size and Subperiod

Table 5 shows the estimations of the ESG portfolios controlled by subperiod as here the portfolios are divided into subperiods split evenly across the full period going from 2006-2013 and 2014-2021.

Table 5 Alphas from difference portfolio (100 firms) controlling for weighting method and subperiods

	Equally Weighted			Market Cap Weighted		
	Full Period	2006 2013	2014 2021	Full Period	2006 2013	2014 2021
USA						
ESG	-0,003	0,000	-0,006	0,002	-0,002	0,005
ENV	-0,001	0,000	-0,002	0,005*	0,007*	0,001
SOC	-0,005**	-0,001	-0,009**	0,001	-0,002	0,005
GOV	0,000	-0,001	0,000	0,003	-0,001	0,007**
Europe						
ESG	-0,005**	-0,004***	-0,005	-0,002*	-0,005**	0,000
ENV	-0,001	-0,005***	0,003*	-0,001	-0,005**	0,004**
SOC	-0,004*	-0,005***	-0,003	0,001	0,000	0,002
GOV	-0,002	-0,002	-0,001	0,003*	0,004	0,001
Emerging^						
ESG	-0,004		-0,004	-0,001		-0,002
ENV	-0,003		-0,003	-0,001		-0,003
SOC	-0,003		-0,004	0,000		0,000
GOV	-0,004*		-0,002	0,002		0,002

*This table shows the alphas of the Carhart (1997) four-factor model over a sample period from 2006 to 2021 including two subperiods on a monthly basis. Regressions are run for each portfolio individually. Including high-low portfolios for each individual category. The high-low portfolio represents the difference portfolios where a long position is taken in the high rated portfolio and a short position in the low rated portfolio. *.**.*** imply a significance level of 10,5 and 1% respectively based on the alphas p-value. ^For the Emerging market the full period is between 2009-2021*

This immediately reports some eye-catching results seeing the evident decrease in the significance of negative abnormal returns in the difference portfolios for Europe between the first and second period, indicating that there has been a decrease in the gap of abnormal returns between the high and low portfolios.

In table 6, the number of firms selected in each portfolio has decreased from 100 to 20. The most noticeable difference is the increasing value and variation of the alphas in the different portfolios this generates. The natural explanation for this is that making less diversified portfolios, i.e., decreasing the number of assets held in the portfolio, results in more volatile outcomes. This is further reflected when higher alpha values like -0,017 for a portfolio can be nonsignificant at all in these portfolios. However, abnormal return values of -0,004 could be significant at the 1% level in the portfolios containing 100 assets. The reason for this is due to the increase in the standard deviation of the portfolios that naturally comes when diversification decreases. Reducing the number of assets held in the various ESG portfolios does not seem to alter the overall results.

Table 6 Alphas from difference portfolios (20 firms), controlling for weighting method and subperiods

	Equally Weighted			Market Cap Weighted		
	Full	2006	2014	Full	2006	2014
	Period	2013	2021	Period	2013	2021
USA						
ESG	-0,004	0,000	-0,009	0,005	-0,001	0,013**
ENV	-0,002	-0,006**	0,002	0,001	-0,003	0,005
SOC	-0,007	0,004	-0,017*	0,003	0,003	0,003
GOV	0,002	-0,001	0,005	0,006*	0,004	0,007
Europe						
ESG	-0,003	-0,007**	0,001	-0,003	-0,006*	-0,001
ENV	-0,001	-0,004	0,002	-0,001	-0,001	-0,002
SOC	-0,014*	-0,011***	-0,017	-0,005*	-0,011***	0,001
GOV	-0,004*	-0,007**	-0,002	0,001	0,000	0,002
Emerging^						
ESG	-0,008		-0,007	0,000		0,003
ENV	-0,005		-0,004	0,000		0,001
SOC	-0,008		-0,013	-0,003		-0,006
GOV	-0,002		-0,004	-0,002		0,003

*This table shows the alphas of the Carhart (1997) four-factor model over a sample period from 2006 to 2021 including two subperiods on a monthly basis. Regressions are run for each portfolio individually. Including high-low portfolios for each individual category. The high-low portfolio represents the difference portfolios where a long position is taken in the high rated portfolio and a short position in the low rated portfolio. *.**.*** imply a significance level of 10,5 and 1% respectively based on the alphas p-value. ^For the Emerging market the full period is between 2009-2021*

The European market continues to show a significant discrepancy between the high and low portfolios in the first subperiod, which deteriorates in the second. Alphas in the emerging markets remain nonsignificant in both weighting methods.

5.4.2 Controlling Using a Best-In-Class Approach

Here the best-in-class approach is implemented, containing an equal number of firms from each sector but keeping the portfolio size the same. Table 7 reaffirms the results for the high minus low portfolios in the U.S that were presented over the entire period in table 2. A slight decrease in the alpha value for the equally weighted ESG portfolio over the full period makes the negative abnormal returns significant here, which is mainly attributed to the second period, which is the same when not applying best-in-class. For the high (low) portfolios in the U.S, both the high rated environmental and social portfolios showed significantly negative alphas.

Table 7 Alphas from difference portfolios (100 firms) using best-in-class approach, controlling for weighting method and subperiods

	Equally Weighted			Market Cap Weighted		
	Full	2006	2014	Full	2006	2014
	Period	2013	2021	Period	2013	2021
USA						
ESG	-0,005**	-0,002	-0,008**	-0,001	-0,006***	0,004
ENV	-0,001	-0,001	-0,003*	0,002	0,002	0,001
SOC	-0,004**	-0,001	-0,008**	0,002	0,001	0,003
GOV	0,001	0,000	0,002	0,004*	0,000	0,008**
Europe						
ESG	-0,003	-0,003*	-0,003	0,002	0,000	0,003
ENV	-0,001	-0,003*	0,001	0,000	-0,004*	0,003
SOC	-0,004*	-0,005**	-0,002	0,001	0,001	0,002
GOV	0,000	-0,001	-0,002	0,000	0,001	0,000
Emerging						
ESG			-0,004			0,001
ENV			-0,001			0,000
SOC			-0,004			0,001
GOV			-0,003			0,001

*This table shows the alphas of the Carhart (1997) four-factor model over a sample period from 2006 to 2021 including two subperiods on a monthly basis. Regressions are run for each portfolio individually. Including high-low portfolios for each individual category. The high-low portfolio represents the difference portfolios where a long position is taken in the high rated portfolio and a short position in the low rated portfolio. *.**.*** imply a significance level of 10,5 and 1% respectively based on the alphas p-value*

Table 11 in the appendix shows that this is suggested for this method as well. In general, for the equally weighted portfolios, the significance and negative values of the portfolios with high-rated assets seem to increase between the periods while their counterparts have a positive

increase in abnormal returns in the same time span. Weighing by market cap gives mostly negative for all portfolios high and low except for the portfolios based on the governance pillar. The significance for the negative alphas is greater in the high portfolios than in the low.

Applying the best-in-class approach results in the same pattern for the European market whereas before, even if the strength of the significancy shrinks somewhat.

Decreasing the methods portfolio size from 100 to 20 for the best-in-class method, as presented in table 8, show no substantial disagreement in the previous results and significance more than an increase in the variance of the alpha values credited to the decrease in diversification. The pattern for the European market remains.

Table 8 Alphas from difference portfolios (20 firms), using best-in-class approach controlling for weighting method and subperiods

	Equally Weighted			Market Cap Weighted		
	Full Period	2006 2013	2014 2021	Full Period	2006 2013	2014 2021
USA						
ESG	-0,005	-0,002	-0,008	0,006*	0,000	0,013**
ENV	-0,001	0,001	-0,003	-0,003	-0,002	-0,003
SOC	-0,008**	0,000	-0,016**	0,003	0,004	0,003
GOV	0,001	-0,001	0,001	0,007**	0,006	0,008*
Europe						
ESG	-0,002	-0,006*	0,003	0,000	-0,003	0,004
ENV	-0,003	-0,007**	0,001	0,000	-0,006*	0,006
SOC	-0,015*	-0,009**	-0,021	-0,002	-0,006*	0,003
GOV	-0,002	-0,003	-0,001	0,006*	0,004	0,008
Emerging						
ESG			-0,004			-0,003
ENV			-0,006			0,000
SOC			-0,016			-0,003
GOV			-0,002			0,006

*This table shows the alphas of the Carhart (1997) four-factor model over a sample period from 2006 to 2021 including two subperiods on a monthly basis. Regressions are run for each portfolio individually. Including high-low portfolios for each individual category. The high-low portfolio represents the difference portfolios where a long position is taken in the high rated portfolio and a short position in the low rated portfolio. *.**.*** imply a significance level of 10,5 and 1% respectively based on the alphas p-value*

5.5 General Discussion of the Results

The three first tables covering all markets and focusing on the initial portfolio construction across the full sample period do not indicate any clear-cut significant themes for any portfolio pillars across regions. The pattern that sticks out is the consistently higher value on the alpha for the portfolios containing low-rated ESG firms compared to their high-rated counterparts. Portfolios constructed using equal weights follow the same pattern when compared to market cap weighted equivalents suggesting that a portfolio construction using equal weights is to be preferred when selecting based on ESG ratings. Returning more positive results from equal weights follows the claims in a recent study by Dorfleitner, Kreuzer and Sparrer (2020)

The result, in general, from portfolios built on market capitalization weights for all markets and over the full period, shows us that there is very little evidence for the support in past literature (see Kempf & Osthoff, 2007; Eccles, Ioannou & Serafeim, 2014) that significant positive abnormal returns can be provided by investing in any combination of high (low) rated ESG portfolios using this weighting method. When splitting the sample period, there does seem to be an improvement in the alphas for the difference portfolios over time using this weighting method. Reporting almost solely positive abnormal returns between 2014-2021, even though most of them are nonsignificant. However, when looking at the tables in the appendix, this approach shows us almost entirely negative alphas for both the high and the low portfolios in all categories except the governance pillar. This underperformance for both portfolios in relation to the market benchmark would, in reality, be even harsher as transaction costs are excluded in these results.

For the U.S. market, there are significant negative returns for the high portfolios in the environmental and social pillars using both weight methods over the full period. When looking at the robustness portfolios, the negative results seem to exclusively be attributed to the second period using equally weighted positions but constantly negative using the other weight method. The same results apply to the best-in-class approach. While not significant, the results favor the low portfolios throughout. There also seems to be a case for significance in the difference portfolio for the social pillar in the recent subperiod.

Looking at the results from the European market, which in the literature has been argued not to feature any substantial positive abnormal returns but rather suggestions of significant negative risk-adjusted returns (Auer & Schuhmacher, 2016). The finding in this paper correlates with those suggestions when looking at the results for the high and high minus low portfolios. The

significant positive abnormal returns for the low portfolios also seem to fall in line with current literature as Bannier, Bofinger and Rock (2019) reported the same results for low portfolios. These results seem to have been more present in the first subperiod than the most recent one, suggesting that this mispricing has somewhat faded.

The European market also reports more noteworthy results. First, from table 3, we know that the difference portfolio in Europe showed significant negative abnormal returns for both the overall ESG and the social portfolios. Then in table 5, controlling for subperiods, the strongly significant negative alphas in Europe for the first period mostly disappear in the second period. This result is in line with the findings of Dorfleitner, Kreuzer and Sparrer (2020), suggesting that while the difference portfolios have had significantly negative anomalous returns in the past, they no longer seem to have it. While this is true, this does not suggest that portfolios containing firms with high ESG ratings in the most recent period have returned better alphas, only in comparison to the low portfolios. When investigating table 9 in the appendix illustrates the same results as table 5. However, for the individual high (low) portfolios, this shows that the most prominent reason for the decrease is due to the low-rated portfolios having significant positive alphas in the first subperiod and performing worse in the second period. This highlights an issue when only considering the high minus low portfolios. For the market cap weighted portfolios, this issue becomes even more apparent as the high minus low portfolios show us a positive alpha, but when investigating their respective high and low portfolio, they both report highly significant negative abnormal returns. This is featured in multiple instances for the developed markets in this study.

Emerging markets seem to have the most robust results as the ESG portfolios consistently show no indication of significant risk-adjusted returns in any categories. The same result holds when checking for robustness based on size, time, and best-in-class approach. The lack of significant results for this market does contradict the findings in the few studies that cover this region, as Friede, Busch and Bassen (2015) suggest that opportunities for abnormal return by ESG investing exist in emerging markets. The results given here could be considered a good starting point for this market. With similar beliefs as Pollard, Sheerwood and Klobus (2018), the sharp growth in ESG rating coverage by firms in these countries, in addition to an increasing sample period for this region. It will undoubtedly lead to more promising and robust conclusions.

Emerging markets and the European markets consisting of multiple countries propose an alternative method to portfolio construction using a similar approach to the best-in-class method. Instead of adjusting for potential bias towards specific industries, it could instead be

altered to control for potential bias towards specific countries in these markets. This paper does not adjust for this bias, meaning that there is a possibility that the high (low) rated portfolios include an overrepresentation of firms in certain countries.

The repeated characteristics shown in the different tables that both the high and low portfolios for some categories and markets have significant adverse risk-adjusted returns brings up the idea of possibly including other portfolios in these types of studies. Some papers using ESG portfolios include much larger sizing in terms of the percentage of the portfolios than in this study. Some even create portfolios of all the sampled companies, making the high (low) portfolios consist of half of all firms (Halbritter & Dorfleitner, 2015). However, the standard approach uses a portfolio construction based on a smaller proportion of the sampled firms, using only the very best and worst rated firms. Creating a third portfolio consisting of more ESG neutral rated firms having ESG scores around the categories mean and then building two additional difference portfolios in the form of high minus average and average minus low. This could introduce an interesting concept for future studies in ESG investing

6. Conclusion

Using a simple and practical approach that most investors could apply, this paper has estimated different sustainable investing strategies by creating ESG portfolios to analyze the relationship between corporations' social and financial performance. This has been done by using one of the literature's more extensive datasets of ESG ratings from Refinitiv, including around 8000 companies between 2006-2021, concentrating on three different regions, the U.S, Europe, and Emerging markets. Previous literature on the topic of ESG investing has been very scarce in emerging markets, making this study one of the first to implement an increasingly researched methodology of ESG portfolios in this region. The outcome of implementing the Carhart (1997) four-factor model indicate that over the full period across the three observed markets, there is a clear absence of significant results pointing towards any consistent positive abnormal returns to be made by executing similar ESG portfolios for an investor. These results contradict findings in previous literature (see Kempf and Osthoff, 2007; Statman & Glushkov, 2009; Eccles, Ioannou & Serafeim 2014). More recent studies, however, have proposed the same lack of substantial positive abnormal returns (see Borgers et al. 2013; Halbritter & Dofleitner 2015). However, there are a few other notable results found in this study.

The perhaps most robust case for consistent results in this essay is found in the market aimed to fill a research gap as the results suggest consistently non-significant abnormal performance using ESG investing in emerging markets. A result that goes against the suggestions by Friede. Busch and Bassen (2015) but should more be seen as an early building block in the research on emerging markets using ESG portfolios. Furthermore, as the steep increase in international firm coverage by ESG rating agencies continues and as the sample period grows. The continued interest in these types of studies seems inevitable for all markets but is bound to shed additional light on less researched markets such as emerging markets

An overall pattern across the observed markets seems to be that the ESG portfolios constructed using equal weights rather than weights based on market capitalization perform better. Generating more positive or at least less negative alpha values for the high (low) portfolios, although they in many cases are nonsignificant. Continuing with equal weights, another pattern found is that low portfolios have superior alphas to high ones.

The most consistent significant results of abnormal returns are found in the European market. Using equal weights for both the low and the difference portfolios shows highly significant

alphas for most ESG categories and are constant even when applying all the different robustness tests. The low portfolios show positive abnormal returns, while the difference portfolios report negative abnormal returns. Results that indicate a potential investment strategy in either the low portfolios or an inverse of the high minus low portfolio (low-minus-high) could generate substantial positive abnormal returns. However, this only holds for the first subperiod between 2006-2013. For the most recent subperiod, the significance of these result has all but disappeared, suggesting that the possible mispricing have been corrected.

Generally, except for the first subperiod in the European market and perhaps the second period in the U.S. using the best-in-class approach, the results of the high minus low portfolios across the regions suggest no significant difference in abnormal returns between investing in an ESG portfolio based on low-rated firms or a portfolio based on high-rated firms. A result that does not reaffirm the "doing well by doing good" prospect in some ESG investing literature. If an investor, however, stands between the choice of investing in a portfolio containing higher or lower ESG-rated firms, perhaps "doing just as well by doing good" is not such a bad deal.

However, the approach of only focusing on the high minus low portfolios, which is commonly used in the literature, disregards one crucial issue. Since these studies usually aim to investigate if there are significant abnormal returns to be made by applying ESG portfolios to an investing strategy, it should be important whether the underlying portfolios of the high minus low portfolios are both positive or both negative. If the high (low) portfolios have significant alphas with the same sign, there, of course, still exists the possibility of abnormal returns by creating a difference portfolio of the two, but it does raise the question of whether this is the most optimal choice or if the investor should seek to combine other ESG portfolio strategies.

To summarize, even though this study is unsuccessful in finding any portfolio strategy with definite significant abnormal returns that holds for the whole period. A big portion of the results in this paper coincide with recent literature reaffirming many of the past conclusions as well as providing new results to build upon. The general presence but inconsistency of significant alphas across the different portfolios suggests that investors seeking to implement investment strategies with the aim of beneficial opportunities should be cautious but curious in the attempts to utilize strategies based on ESG portfolios.

The conclusions of this thesis do propose some alternative methods to consider. The recurring evidence across markets and subperiods suggesting negative abnormal returns in both the high and low portfolios, mainly using market cap weights, advocates motivation for including

another portfolio selection of ESG-rated firms. An approach that includes firms not only ranked at the very top or bottom but a portfolio consisting of firms with more neutral ESG ratings around the mean. Applying difference portfolios between the top and middle rated as well as the bottom and middle rated portfolios could lead to a new perspective on portfolio construction in ESG investments. However, that will be left for future studies to investigate.

References

Auer, B. R., & Schuhmacher, F. (2016). Do socially (ir)responsible investments pay? New evidence from international ESG data, *The Quarterly Review of Economics and Finance*, vol. 59, pp.51-62

Bannier, C. E., Bofinger, Y., & Rock, B. (2019). Doing safe by doing good: ESG investing and corporate social responsibility in the U.S. and Europe. CFS Working Paper Series, no. 621, Goethe University, Center for Financial Studies.

Berg, F., Kölbel, J., & Rigobon, R. (2019). Aggregate Confusion: The Divergence of ESG Ratings. *Forthcoming Review of Finance.*

Bhagat, S. & Bolton, B., (2008). Corporate governance and firm performance, *Journal of Corporate Finance*, vol. 3, pp. 257-273

Bloomberg (2021) Europe Expands Dominance in ESG Market After New Rules Enforced, Available online: https://www.bloomberg.com/news/articles/2021-11-11/europe-expands-dominance-in-esg-market-after-new-rules-enforced [Accessed 17 August 2022]

Borgers, A., Derwall, J., Koedijk, K., & Horst, J, T. (2013). Stakeholder Relations and Stock Returns: On Errors in Investors' Expectations and Learning, *Journal of Empirical Finance*, vol. 22, pp.159–175

Bowen, H. R. (1953). Social responsibility of the businessman. New York: Harper & Row

Busch, T., & Hoffmann, V. H. (2011). How Hot is Your Bottom Line? Linking Carbon and Financial Performance. *Business and Society*, vol. 50, no. 3, pp.233-265

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

Cheng, B., Ioannou, I., & Serafeim, G. (2013). Corporate social responsibility and access to finance. *Strategic Management Journal*, vol.35, no.1, pp.1-23

Chong, J., & Phillips, M. (2013). Portfolio Size Revisited *Journal of wealth management* vol.15, no.4, pp.49-60

Dobers, P., & Halme, M., (2009). Corporate social responsibility and developing countries. Corporate Social Responsibility and Environmental Management, vol.16, no.5, pp.237-249

Dorfleitner, G., Kreuzer, C., & Sparrer, C. (2020). ESG controversies and controversial ESG: about silent saints and small sinners. *Journal of Asset Management, vol.21*, pp.393-412

Eccles, R. G., Ioannou, I., & Serafeim, G. (2014). The Impact of Corporate Sustainability on Organizational Processes and Performance. *Management Science,* vol.60, no.11, pp.2835–2857.

Eccles, R. G., Lee, L., & Stroehle, J.C. (2020). The social Orgins of ESG: An Analysis of Innovest and KLD. *Organization & Environment,* vol.33, no.4, pp.575-596.

Endrikat, J., Guenther, E., & Hoppe, H. (2014). Making sense of conflicting empirical findings: A meta-analytic review of the relationship between corporate environmental and financial performance. *European Management Journal, vol.*32, no.5, pp.735-751

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. (2004). The Capital Asset Pricing Model: *Theory and Evidence. Journal of Economic Perspectives*, vol.18, no.3, pp.25-46

Fama, E. F., & French, K. R. (2012). Size, value, and momentum in international stock returns. *Journal of Financial Economics*. vol.105, no.3, pp.457-472

Fama, E. F, & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, vol.116, no.1, pp.1-22

Friede, G., Busch, T., & Bassen, A. (2015). ESG and financial performance: aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment*, vol. 5, no. 4, pp.210-233

Garcia, A. S., Mendes-Da-Silva, W., & Orsato, R. J. (2017) Sensitive industries produce better ESG performance: Evidence from emerging markets. *Journal of Cleaner Production*, vol.150, pp.135-147

Gompers, P., Ishii, J., & Metrick, A. (2003). Corporate governance and equity prices. *The Quarterly Journal of Economics*, vol.118, no.1, pp.107–155

GSIR (2021) Global Sustainable Investment Review 2020, Available online: http://www.gsi-alliance.org/wp-content/uploads/2021/08/GSIR-20201.pdf [Accessed 18 August 2022]

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

Hong, H., & Kacperczyk, M. (2009) The price of sin: The effects of social norms on markets. *Journal of Financial Economics*, vol.93, no.1, pp.15-36

Jiang, G. J., Yao, T., & Yu, T. (2007). Do mutual funds time the market? Evidence from portfolio holding. *Journal of Financial Economics*, vol.86, no.3, pp.724-758.

Jegadeesh, N., & Titman, S. (1993) Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. The Journal of Finance, vol.48, no.1, pp.65-91

Kempf, A., & Osthoff, P. (2007). The effect of socially responsible investing on portfolio performance. *European Financial Management*, vol.13, no.5, pp.908–922.

Litner, J. (1965). The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *The Review of Economics and Statistics*, vol.47, no.1, pp.13-37

Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, vol.7, no.1, pp.77-91

Markowitz, H. M. (1959). Portfolio Selection: Efficient Diversification of Investments. Yale University Press.

Morningstar (2021). Sustainable Funds U.S. Landscape Report 2021, Available online: https://www.morningstar.com/lp/sustainable-funds-landscape-report [Accessed 1 August 2022]

MSCI (2022). MSCI Emerging Markets Index (USD), Available online: https://www.msci.com/documents/10199/c0db0a48-01f2-4ba9-ad01-226fd5678111 [Accessed 18 August 2022]

Orsato, J. Renato., Garcia, A., Mendes-Da-Silva, W., Simonetti, R., & Monzoni, M. (2015) Sustainability indexes: why join in? A study of the 'Corporate Sustainability Index (ISE)' in Brazil

Pedersen, L. H., Fitzgibbons, S., & Pomorski, L. (2021) Responsible investing: The ESG-efficient frontier. Journal of Financial Economics, vol.142, no.2, pp.572-597

PRI (2022) Principles for Responsible Investments Annual Report 2021. Available online: https://www.unpri.org/annual-report-2021 [Accessed 18 August 2022]

Pollard, J., Sherwood, M. W., & Klobus, R. G. (2018) Establishing ESG as Risk Premia. *Journal of Investment Management*, vol.16, no.1, pp.1-12

Sahin, Ö., Bax, K., Czado., & Paterlini, S. (2022) Environmental, Social, Governance scores and the Missing pillar -Why does missing information matter? Corporate Social Responsibility and Environmental Management vol. pp.1-17

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

Statman, M., & Glushkov, D. (2009). The wages of social responsibility. *Financial Analysts Journal*, vol.65, no.4, pp.33–46.

Refinitiv (2022a). Environmental, Social and Governance (ESG) scores from Refinitiv, Available online: https://www.refinitiv.com/content/dam/marketing/en_us/documents/methodology/refinitiv-esg-scores-methodology.pdf [Accessed 6 August 2022]

Refinitiv (2022b). The Refinitiv Business Classification Methodology, Available online: https://www.refinitiv.com/content/dam/marketing/en_us/documents/methodology/trbc-business-classifcation-methodology.pdf [Accessed 18 August]

Reuters (2022). Global green finance rises over 100 fold in the past decade -study, Available online: https://www.reuters.com/business/sustainable-business/global-markets-greenfinance-graphics-2022-03-31/ [Accessed 1 August 2022]

UNEP FI (2005) UNEP FI 2005 Overview, Available online: https://www.unepfi.org/fileadmin/documents/unepfi_overview_2005.pdf [Accessed 18 August 2022]

Whittaker, J. (2011). The evolution of environmentally responsible investment: An Adam Smith perspective. Ecological Economics, vol.71, pp.33-41

Appendix

Table 9 Alphas for high and low portfolios (100 firms) controlling for weighting method and subperiods

		Equally Weighted			Market Cap Weighted		
		Full	2006	2014	Full	2006	2014
U.S		Period	2013	2021	Period	2013	2021
ESG							
	High	-0,001	-0,001	-0,002	-0,002***	-0,003***	-0,002*
	Low	0,002	0,000	0,004	-0,004**	-0,001	-0,007*
ENV							
	High	-0,002***	0,000	-0,004***	-0,003***	-0,002**	-0,004***
	Low	-0,001	0,000	-0,002	-0,008***	-0,009**	-0,005***
SOC							
	High	-0,002**	0,000	-0,003**	-0,002***	-0,003***	-0,001
	Low	0,003	0,000	0,006*	-0,003*	-0,001	-0,006*
GOV							
	High	0,001	0,000	0,001	0,000	-0,002	0,003*
	Low	0,001	0,001	0,001	-0,002	-0,001	-0,004
Euro							
ESG							
	High	-0,001	-0,001	-0,001	-0,003***	-0,004***	-0,003***
	Low	0,004**	0,004**	0,004	-0,001	0,001	-0,003*
ENV							
	High	0,000	-0,001	0,000	-0,002***	-0,004***	-0,001
	Low	0,001	0,004**	-0,003**	-0,002	0,001	-0,005***
SOC							
	High	0,000	0,000	0,000	-0,003***	-0,003***	-0,002**
	Low	0,004**	0,005***	0,003	-0,004**	-0,003	-0,004*
GOV							
	High	-0,001	-0,001	-0,001	0,000	0,000	0,000
	Low	0,001	0,002	0,001	-0,003**	-0,005	-0,002
EM^							
ESG	High	-0,001		0,000	-0,002		-0,001
	Low	0,003		0,004	-0,001		0,001
ENV							
	High	-0,002		0,000	-0,002		0,000
	Low	0,001		0,002	-0,001		0,002
SOC							
	High	-0,001		0,001	-0,001		0,001
	Low	0,003		0,005	-0,001		0,001
GOV							
	High	-0,002*		0,000	-0,002		0,001
	Low	0,001		0,002	-0,004		-0,002

*This table shows the alphas of the Carhart (1997) four-factor model over a sample period from 2006 to 2021 including two subperiods on a monthly basis. Regressions are run for each portfolio individually. The high (low) portfolios include the 100 companies with the highest (lowest) rating overall and for each individual category. *.**.*** imply a significance level of 10,5 and 1% respectively based on the alphas p-value. ^ ^For the Emerging market the full period is between 2009-2021*

Table 10 Alphas for high and low portfolios (20 firms) controlling for weighting method and subperiods

		Equally Weighted			Market Cap Weighted		
		Full	2006	2014	Full	2006	2014
U.S.		Period	2013	2021	Period	2013	2021
ESG							
	High	0,001	0,001	0,000	0,001	-0,001	0,003
	Low	0,005	0,001	0,009	-0,005	0,000	-0,010**
ENV							
	High	-0,003***	-0,004**	-0,003*	-0,004***	-0,006***	-0,001
	Low	-0,002	0,003	-0,004	-0,005*	-0,003	-0,006
SOC							
	High	0,000	0,001	-0,001	0,000	0,000	0,001
	Low	0,007	-0,003	0,016*	-0,003	-0,004	-0,002
GOV							
	High	0,003*	0,002	0,004	0,003*	0,000	0,004*
	Low	0,001	0,003	-0,001	-0,003	-0,004	-0,003
Euro							
ESG							
	High	-0,001	-0,001	-0,001	-0,003***	-0,004	-0,002
	Low	0,002	0,005*	-0,002	0,000	0,001	-0,001
ENV							
	High	0,001	0,000	0,001	-0,002	-0,004	0,000
	Low	0,001	0,004	-0,001	0,000	-0,003	0,002
SOC							
	High	-0,002	-0,002	-0,001	-0,003***	-0,003	-0,003*
	Low	0,013	0,009***	0,016	0,002	0,008	-0,004
GOV							
	High	-0,001	-0,001	-0,001	-0,002	-0,003	-0,002
	Low	0,003*	0,005**	0,001	-0,004	-0,003	-0,003
EM^							
ESG							
	High	-0,001		0,001	0,000		0,002
	Low	0,007		0,008	0,000		-0,001
ENV							
	High	-0,003		-0,001	-0,002		0,001
	Low	0,002		0,003	-0,002		0,000
SOC							
	High	-0,001		0,000	-0,003		-0,004
	Low	0,007		0,013	0,000		0,002
GOV							
	High	0,001		0,000	-0,002		0,000
	Low	0,003		0,004	0,000		-0,003

*This table shows the alphas of the Carhart (1997) four-factor model over a sample period from 2006 to 2021 including two subperiods on a monthly basis. Regressions are run for each portfolio individually. The high (low) portfolios include the 20 companies with the highest (lowest) rating overall and for each individual category. *.**.*** imply a significance level of 10,5 and 1% respectively based on the alphas p-value ^For the Emerging market the full period is between 2009-2021.*

Table 11 Alphas for high and low portfolios using the best-in-class approach (100 firms) controlling for weighting method and subperiods

		Equally Weighted			Market Cap Weighted		
		Full	2006	2014	Full	2006	2014
U.S		Period	2013	2021	Period	2013	2021
ESG							
	High	-0,002**	-0,001	-0,003**	-0,004***	-0,005***	-0,002**
	Low	0,003	0,001	0,006	-0,003	0,001	-0,006
ENV							
	High	-0,002***	-0,001	-0,004***	-0,003***	-0,004***	-0,002*
	Low	-0,001	-0,001	-0,001*	-0,005	-0,006	-0,004*
SOC							
	High	-0,002**	-0,001	-0,003***	-0,003***	-0,003***	-0,002**
	Low	0,003	0,001	0,005	-0,005	-0,004	-0,005*
GOV							
	High	0,000	0,001	0,000	0,002	0,001	0,002
	Low	-0,001	0,001	-0,002	-0,002	0,001	-0,006**
Euro							
ESG							
	High	0,000	0,000	0,000	-0,003***	-0,003***	-0,002***
	Low	0,003	0,003**	0,003	-0,004**	-0,003	-0,005**
ENV							
	High	0,000	0,000	0,000	-0,003***	-0,004***	-0,002**
	Low	0,001	0,003*	-0,001	-0,003*	0,000	-0,005***
SOC							
	High	0,000	0,000	0,000	-0,003***	-0,003***	-0,002***
	Low	0,003*	0,005***	0,002	-0,004**	-0,003	-0,004*
GOV							
	High	0,000	0,001	-0,002*	-0,003***	-0,003***	-0,003***
	Low	0,001	0,002**	0,000	-0,003**	-0,004	-0,002
EM							
ESG							
	High			-0,001			0,000
	Low			0,003			-0,001
ENV							
	High			0,000			0,000
	Low			0,001			-0,001
SOC							
	High			0,001			0,001
	Low			0,005			0,000
GOV							
	High			-0,001			0,000
	Low			0,003			-0,002

*This table shows the alphas of the Carhart (1997) four-factor model over a sample period from 2006 to 2021 including two subperiods on a monthly basis. Regressions are run for each portfolio individually. The high (low) portfolios include the 100 companies with the highest (lowest) rating overall and for each individual category selected based on a best-in-class approach. *.**.*** imply a significance level of 10,5 and 1% respectively based on the alphas p-value.*

Table 12 Alphas for high and low portfolios using best-in-class approach (20 firms) controlling for weighting method and subperiods

		Equally Weighted			Market Cap Weighted		
		Full	2006	2014	Full	2006	2014
U.S		Period	2013	2021	Period	2013	2021
ESG							
	High	0,000	0,000	-0,001	0,002	0,000	0,004*
	Low	0,004	0,002	0,007	-0,004	0,000	-0,009**
ENV							
	High	-0,002	-0,001	-0,002	-0,004***	-0,003*	-0,004**
	Low	-0,001	-0,003	0,001	-0,001	-0,001	-0,001
SOC							
	High	-0,002	0,002	-0,005***	0,000	0,002	-0,001
	Low	0,006*	0,001	0,011	-0,003	-0,002	-0,005
GOV							
	High	0,001	-0,001	0,003	0,003**	0,001	0,005**
	Low	0,001	0,000	0,002	-0,004	-0,005	-0,003
Euro							
ESG							
	High	-0,001	-0,002	0,000	-0,003**	-0,003**	-0,002
	Low	0,001	0,004	-0,003	-0,003	0,000	-0,005
ENV							
	High	0,000	-0,002	0,001	-0,001	-0,003**	0,002
	Low	0,002	0,005	0,000	-0,001	0,003	-0,005
SOC							
	High	-0,002	-0,001	-0,002	-0,003**	-0,002	-0,003*
	Low	0,014*	0,008*	0,019	-0,001	0,004	-0,006
GOV							
	High	-0,002	0,000	-0,003*	-0,003*	-0,003**	-0,001
	Low	0,000	0,002	-0,003	-0,009***	-0,007	-0,009**
EM							
ESG							
	High			0,000			-0,002
	Low			0,004			0,001
ENV							
	High			-0,002			0,001
	Low			0,004			0,001
SOC							
	High			-0,001			-0,002
	Low			0,015			0,001
GOV							
	High			0,001			-0,002
	Low			0,003			-0,008*

*This table shows the alphas of the Carhart (1997) four-factor model over a sample period from 2006 to 2021 including two subperiods on a monthly basis. Regressions are run for each portfolio individually. The high (low) portfolios include the 20 companies with the highest (lowest) rating overall and for each individual category selected based on a best-in-class approach. *.**.*** imply a significance level of 10,5 and 1% respectively based on the alphas p-value.*