Stockholm School of Economics

Master Thesis

The ESG Dilemma

ABSTRACT

Investing with the purpose of doing good has been gaining a lot of momentum in the past few years. Leading to increased investments in exchange traded products, mutual funds and equities with higher levels of ESG score both in developed and developing countries. However, it is still unclear whether or not an investor might be able to achieve higher returns in one of these markets and if the individual pillar scores in the ESG measurement could have an effect on the actual returns. There are two main purposes of this study. First, we want to evaluate if a portfolio of highly rated ESG stocks from developing markets will deliver a risk-adjusted excess return over a portfolio of low rated ESG stocks from developed markets. Second, we evaluate if changes in a firm's one-year lagged ESG pillar scores which are environmental, social and governance, affects the firm's stock returns differently in developed and developing markets. We find that developed market stocks have a higher risk-adjusted return over developing market stocks when ESG risk is controlled for. Furthermore, a portfolio long in high ESG-rated developing market stocks and short in low ESG-rated developed market stocks return a significant negative alpha.

Keywords:
Sustainable Investing, SRI, ESG

Authors:
Marcus Galli (50499)
Lamin K. Darboe (41676)

Tutor:
Riccardo Sabbatucci

Examiner:

Master Thesis
Master Program in Finance
Stockholm School of Economics

Table of contents

1. Introduction **2**

2. The Evolution of SRI and ESG as measurements **4**

3. Literature Review **6**

3.1 Opposing literature on risk-return tradeoff of ESG investing 6

3.2 Behavioral aspects of ESG investing 7

3.3 ESG investing in developed and developing markets 10

3.4 Relationship between firm performance and ESG 11

4. Hypotheses **13**

5. Empirical Work **14**

5.1 Data 14

5.2 Empirical Method 17

5.2.1 Panel Regression 17

5.2.2 Carhart (1997) Four Factor Model 19

5.2.2.1 Portfolio Construction 19

5.3 Results 22

5.3.1 Panel Regression Results 22

5.3.2 Carhart (1997) Four Factor Regression 25

6. Conclusion **27**

6.1 Main findings 27

6.2 Discussion 27

6.2.1 Implications for investors 27

6.2.2 Comparison to other findings in literature 28

6.2.3 Limitations 28

6.3 Contribution to academia 29

6.4 Future research 30

7. References **31**

7.1 Articles 31

7.2 Links and websites 33

8. Appendix **34**

8.1 Refinitiv ESG Materiality Matrix 34

8.2 Pillar Scoring Matrix 35

8.3 Panel regression with robust SE 36

8.4 Equal-weighted zero-cost portfolio return (%) 37

8.5 Value-weighted zero-cost portfolio return (%) 37

8.6 Equal-weighted long portfolio return (%) 38

8.7 Value-weighted long portfolio return (%) 38

8.8 Equal-weighted short portfolio return (%) 39

8.9 Value-weighted short portfolio return (%) 39

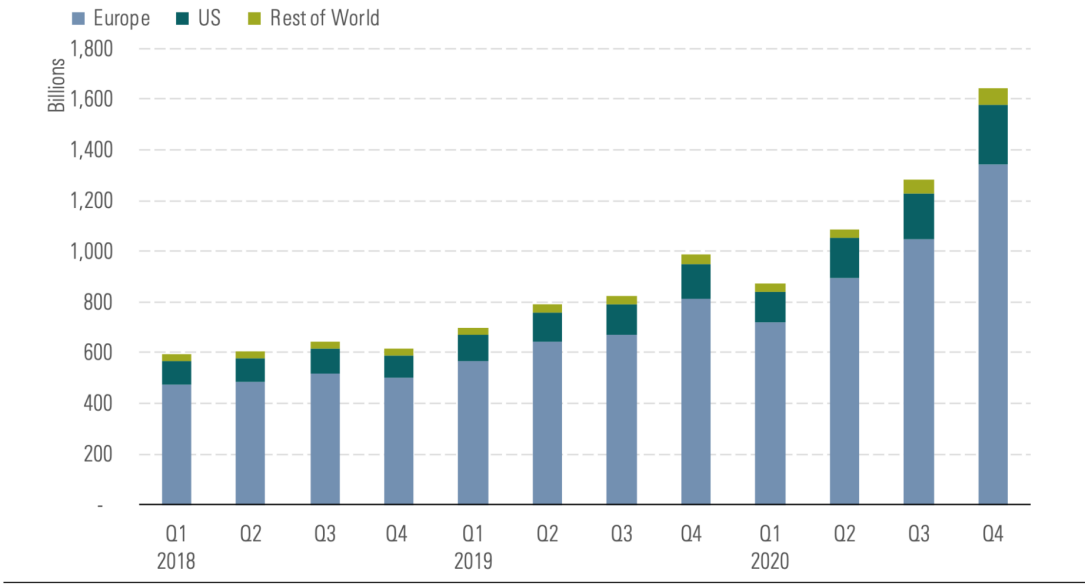
1. Introduction

Investing in a socially responsible manner, often referred to as "socially responsible investment" (SRI) has been growing significantly over the last few years. This growth can both be explained by an increasing investor demand or exclusion there-off from the buy side, but also as an action from companies on the sell side. The incorporation of these types of policies in investment and business practices can be explained by multiple factors. Generally, it is a way for companies to do good and neutralize the footprint that they have on our planet. In some cases, it is a possibility to improve the image or performance of either the company itself or one of its brands, usually with hopes of better future financial gains (Velte, 2017). Lastly, there has been an increased pressure on countries and politicians to incorporate these types of policies into laws and regulations, which to an extent forces companies to become increasingly ESG compatible.

Although ESG investing is ethically and morally soothing for an investor, an equally important factor in this decision is the potential difference in risk adjusted return, return per quantity of risk, compared to a non ESG portfolio. Previous studies show that individuals who engage in this type of investing consider both the effect that the investment will have on their social status, how the individual will be perceived by others in their surroundings, and the potential increase in returns that the decision might have (Anderson and Robinson, 2019). Interestingly, the study shows that expectations of returns have a big impact on the decision making of the individual. If the investor expects that the returns of the ESG portfolio will exceed that of the non ESG portfolio, then that person is more likely to choose the ESG portfolio and vice versa.

According to *Morningstar global ESG flows report for the fourth quarter 2020*, inflows to global sustainable funds had increased by 88% in Q4 2020 to a total amount of 152,3 billion USD. A majority of these inflows, approximately 80%, were accounted for by Europe, compared to the US who accounted for 13,4%. In addition, Asia ex-Japan accounted for 3,3%, Australia/New Zealand 0,8%, Japan 3,7% and Canada 0,8%. Looking at these numbers, Europe would be considered the biggest market for SRI investing, however, as this trend continues to influence society, businesses and politics, existing markets and new markets are likely to continue growing.

Figure 1: Quarterly Global Sustainable Fund Assets (USD Billion)



Source: Morningstar Direct, Manager Research. Data as of December 2020.

The data highlights a clear skewness in flows towards developed markets for ESG investing. This made us question whether or not ESG is priced to the same extent in developing countries given the huge inflows which have been seen in developed markets for the past years. Thus, in this study, we want to evaluate the investment opportunities presented by high-ESG rated stocks in emerging countries. This research could in turn, given a finding of differences between returns, shed light on the opportunities of developing markets and thus increase the fund flows to the developing equity markets.

2. The Evolution of SRI and ESG as measurements

Financial theory highlights an important issue that usually arises during discussions about sustainable investing. According to Berry and Junkus (2011), there is no real consensus as to what Socially Responsible Investing, or SRI, means to an investor. The action itself of investing responsibly, could be traced back to different historical movements, often with religious roots. However, the action of investing with this approach has been gaining increasing momentum over the last few years. According to the *Eurosif European SRI Study 2018*, the fund flows in Europe have been growing at an annual growth rate (CAGR) of about 20% from 2009 to 2017. Affecting not only the equity market, but also the bond market which is experiencing an enormous green wave. Putting even more pressure on having both an accurate measurement and definition of what is to be considered SRI and not.

Both of these studies highlight the fact that a vague definition has become an ominous problem for the movement itself. Funds and investment vehicles that might consider themselves to be sustainable might in fact be the total opposite. *Eurosif* has therefore reached a consensus for the definition of SRI which in turn represents the European level and view. According to *Eurosif*, SRI should be defined as follows:

> *"Sustainable and responsible investment (SRI) is a long-term oriented investment approach which integrates ESG factors in the research, analysis and selection process of securities within an investment portfolio. It combines fundamental analysis and engagement with an evaluation of ESG factors in order to better capture long term returns for investors, and to benefit society by influencing the behavior of companies."*

In turn, this highlights another fundamental keyword for sustainable investing which is usually referred to as ESG. The abbreviation is short for Environmental, Social, Governance and is often used in the decision-making process by investors and fund managers of different kinds when evaluating different investments (Dorfleitner et al, 2015). ESG generally takes the form of a score, either from a scale of 0-100 or A+ to D, which is a way for investors to quickly get an idea of how the company is performing for the three pillars included in ESG (Refinitiv, 2021). These data points are put together and summarized by different actors such as Bloomberg or Refinitiv which in turn can be accessed through their respective terminal, Bloomberg or Eikon and is generally updated on a yearly basis.

On the contrary, Escrig-Olmedo et al (2019) highlights an important problem regarding ESG rating agencies and their ability to fully integrate sustainability principles. The ESG rating industry has been through multiple mergers and acquisitions in recent years, making the sector more professional and connected to financial companies. However, ESG rating companies have a commercial character since they market and sell their products, possibly making them more biased towards sustainability. Which in turn could have an impact on their assessment of businesses. Also, when evaluating businesses based on sustainability, the authors highlight that rating institutes should try to further integrate the sustainability criteria in their rankings. For example including methods such as footprint analysis which provides a life-cycle approach that would give a better assessment and a more comprehensive sustainability measurement.

3. Literature Review

The literature review section is divided into four different parts which are supposed to give a good overview and understanding of SRI and ESG investing. First, we discuss the risk-return tradeoff of ESG investing. Second, the behavioral aspects that can have an effect when making these types of investment decisions. Third, differences between developed and developing markets. Lastly, the relationship between firm performance and the given ESG rating.

3.1 Opposing literature on risk-return tradeoff of ESG investing

In addition to the definition of ESG, a highly relevant topic is whether or not there are tradeoffs of choosing an ESG portfolio compared to a regular portfolio. Vlassis (2020) provides helpful insights on this topic with one of the to date longest time series studies for ESG. The paper shows that portfolios consisting of companies with high ESG ratings do not suffer in terms of risk-adjusted returns, when compared to companies considered not to be ESG. In some cases, these highly rated companies even outperform the market.

Furthermore, the paper highlights an interesting aspect of the ESG scoring, mainly that firm size, in terms of market cap, seems to play an important role when explaining the ESG score of a firm, also noted by Drempetic et al (2019). Their study is based on Thomson Reuters ASSET4 ESG rating and finds significant correlation between firm size and ESG rating. The main reason for this is due to the fact that larger firms provide more data for ESG databases, making it easier to give them a specific rating. However, this does not necessarily mean that larger firms are more sustainable than smaller firms, only that data of larger firms, and thus their scoring is more comprehensive.

Although some studies show that an ESG oriented portfolio could generate excess returns compared to conventional ones, there are some studies with opposing results. Halbritter and Dorfleitner (2014) use several ESG rating agencies, Bloomberg, Thomson Reuters ASSET4 and KLD to construct a high-low difference portfolio of ESG companies from 2002 to 2011. By using three data providers they get a more comprehensive dataset compared to previous studies which generally only use one provider. Their results indicate that the ESG portfolios do not generate significant excess returns compared to traditional ones, even when using

different cut-offs and weightings. Also, pursuing a best-in-class ESG portfolio provides similar results, not generating any abnormal returns.

Bannier et al (2019), based on Thomson Reuters ASSET4 data from 2003 to 2017, finds that a portfolio long in stocks with highest ESG scores and short in those with lowest cores yields a significantly negative abnormal returns. Interestingly, this is caused by the strong positive return of firms with lowest ESG activity.

Fatemi et al (2018) provides additional insight regarding disclosure and reporting of ESG and in turn the effects that it might have on the valuation of a firm. Their study shows that high ESG disclosure could be interpreted by the market as an attempt by the firm to justify an overinvestment in sustainability. In turn, weakening the positive valuation effects of any previous ESG strengths.

3.2 Behavioral aspects of ESG investing

When choosing between a sustainable or traditional investment, investors generally face two dilemmas, one of them being excess returns which we have already discussed in previous sections. The second dilemma regards behavioral aspects including both our own perspective and the perspective and thoughts of others. Riedl and Smeets (2017) studied why investors choose to invest in and hold socially responsible mutual funds. Their study shows that, when investing in a sustainable manner, social signaling as well as social preferences play an important role when making these types of investment decisions. Thus, before making a choice, you take into consideration the thoughts that other people would have of your decision. Also, if an investor for some reason were to believe that the returns of the specific sustainable strategy were to provide lower returns, then he or she is less likely to make that decision.

Hong and Kacperczyk (2009) add on to this discussion by researching “sin” stocks, companies involved in producing alcohol, tobacco and gambling. They find that these stocks generally have higher returns but are neglected by bigger institutions and investors due to norm constraints. Making it harder for these companies to obtain debt financing although providing this type of debt usually would provide good returns for the lending party. Proving that investors sometimes value norms and social preferences to a greater extent than actual returns. Hence, social norms and preferences both affect smaller retail investors but also

bigger institutional investors in a similar way. Their paper raises an important aspect of sustainable investing, mainly the fact that investors are willing to, even though they know that returns could be better for these “sin” stocks, pay a financial cost to refrain from them. This finding however only affects the institutions or individuals that are bound to invest according to ESG guidelines, not investors that should generally be perceived as natural arbitrageurs such as hedge funds.

Individual experiences can also have an impact on the way investors choose to allocate their portfolio. For example, if an individual were to be exposed to a natural disaster such as a heatwave, then he or she is more likely to make environmentally responsible investments after that event (Anderson and Robinson, 2019). The reason for this increased weight into sustainable investments is dependent on the probability weighting that the individual places on similar events occurring in the future. Having gone through for example a heatwave, individuals tend to overweight the probability for future climate disasters which in the end makes them tilt their investments into sustainable alternatives. This suggests that behavioral channels and mechanisms could help increase the understanding for portfolio choices made by individuals who have undergone extreme weather events. These persons who overweight future environmental disasters also have a tendency to overpay for investments that are focused on ESG and also perceive themselves to be much better at recycling than their neighbors. Once again highlighting the two important aspects of sustainable investing in terms of returns and social dilemmas.

Petersen et al (2020) decided to create a framework for the problem that investors face in terms of preferences and ESG. First, they divide investors into three different types, Type-U (ESG-unaware) which only want to maximize their mean variance utility, Type-A (ESG-aware) which have mean-variance preferences but use ESG scores to adjust their view on risk and return, and lastly Type-M (ESG-motivated) who have preferences for high ESG scores searching for portfolios with low risk, high expected return and high average ESG scores. By combining mean-variance frontiers and sharpe ratios derived from specific ESG scores, they manage to create a two-dimensional model which shows the mean-variance ESG frontier as shown below.

Figure 2: ESG-efficient frontier

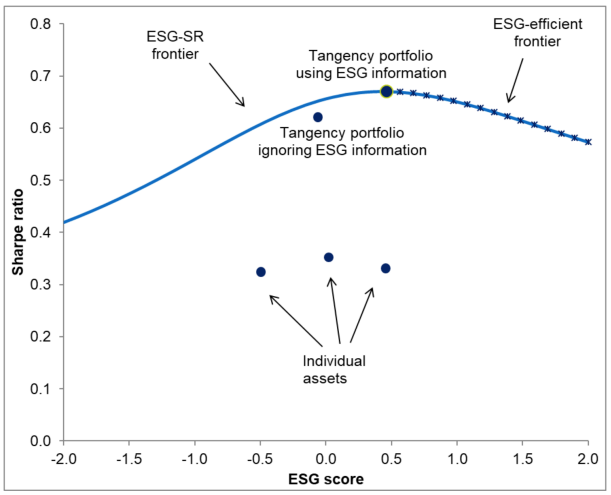
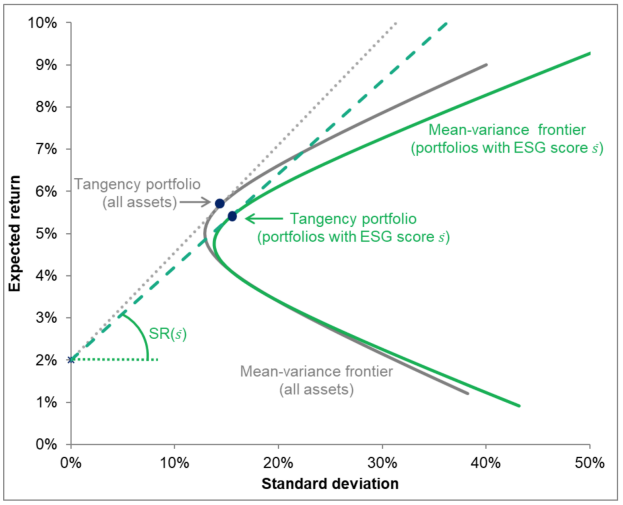


Figure 3: Mean-variance frontiers for all assets and portfolios with certain ESG score



In Figure 2, the unaware investor would choose a portfolio below the frontier since they are unaware of the ESG score of the security and only want to maximize their own mean-variance utility. The motivated investor would choose a portfolio to the right of the tangency portfolio since they want a higher ESG score. Lastly, the aware investor would choose the tangency portfolio because of their mean-variance preferences in combination with ESG scores. These models are important since they help explain the differences in preferences between investors and thus their behavior regarding sustainable investments. In turn, creating a better understanding for the different parameters, as well as setting up a rigorous framework for this process.

3.3 ESG investing in developed and developing markets

Even though sustainable investing has been gaining popularity in recent years, there are still vast differences in the extent to which it is performed and incorporated across countries and especially between developing and developed markets. Even though this incorporation is advancing fast in both markets, the advancements in developing markets might be considered even more crucial. This is because emerging markets are likely to achieve high growth during this century. However, some of the countries that were to be considered developing, are still tackling problems in terms of poverty, urbanization, corruption and pollution. Making it difficult to incorporate some of these sustainable initiatives.

Odell and Ali (2016) have studied ESG investing in emerging and frontier markets to explore any challenges and opportunities that these markets might face when transitioning into sustainability. They highlight that developing markets have been slow to integrate sustainability into their investment process for reasons similar to those mentioned above and that the subject itself is more important for developing than developed markets due to future growth. These evolving markets, although facing short term difficulties, might propose good investment opportunities for investors that are long-term. According to Odell and Ali (2016), from experience, management teams that incorporate these types of issues, are likely to achieve good financial returns over time. This is because they are able to grow faster, reduce cost of capital and manage risks better than peer companies. Companies that integrate sustainability are also easier to differentiate which in turn creates a better competitive advantage and creates value over time. Also avoiding companies that might have risks in the business model that might not be visible from the perspective of the market. In combination with active ownership, these investments can also be a way for fund managers to drive alpha generation.

In addition to this, there has also been studies performed in emerging markets regarding the performance of ESG in different types of industries. Garcia et al (2017) explore this topic by combining Thomson Reuters data with BRICS which includes Brazil, Russia, India, China and South Africa. They discover that companies performing their operations in “sensitive” industries, which is defined as companies subject to “systematic social taboos, moral debates, and political pressures and those that are more likely to cause social and environmental damage”, have superior environmental performance compared to firms in non-sensitive

industries. The authors also highlight that companies classified as “sensitive” do not necessarily need to have a low ESG performance. The risk might be higher, but these companies usually have extensive internal and external control systems and usually manage to achieve high financial performance. Thus, these companies benefit from higher performance in ESG. Also indicating that there is an expectation that firms with a high ESG-rating are able to achieve better performance and in turn higher valuations.

3.4 Relationship between firm performance and ESG

It is clear that fund flows have increased dramatically for sustainable investments during the last few years. However, it is still highly relevant to understand whether or not the flows are based on the belief that a higher ESG rating would indicate better performance for a company or if it is done for moral or social reasons. Hartzmark and Sussman (2019) find that an investor’s heuristic can influence the expectations of future returns for a company. In turn making investment decisions that are truly based on the ESG rating itself and a belief that this in turn could have a positive effect on the financial performance of the company. The study however finds evidence that funds with high rankings or many “globes” , which is the scale that Morningstar uses, does not indicate higher future performance. Thus, the belief that investors have on these highly rated companies does not seem to be correct.

Margolis et al (2009) studied corporate social performance (CSP) and its relation to corporate financial performance (CFP). If doing good would result in doing well, then companies might be inclined to start doing well in terms of addressing social problems, or any bad conduct from the company itself. Even though this type of work might be associated with costs, if there is a positive link between CSP and CFP, companies would still be willing to pursue this work since it could increase their results as well as being a way for the firm to contribute to societal well-being. The results of Margolis et al (2009) study indicates that there is a small positive relationship between corporate social performance and corporate financial performance. However, the results, even if they are statistically significant, are still quite small. For example, corporate social performance only accounts for about 2,23% of the variance in corporate financial performance. Thus, the effect of CSP is relatively low. Also, it is important to note that when doing bad and being discovered, the effect is bigger on financial performance than that of doing good. Which can be the results of behavioral effects where humans tend to weigh negative events to a greater extent compared to positive events.

A similar study was performed in Germany which instead looked at ESG and its impact on financial performance Velte (2017), hence including more factors, environmental and governance, than the study performed by Margolis et al (2009). The paper evaluates any possible links between ESG scores and financial performance with data gathered from Thomson Reuters ASSET4 as well as accounting data such as return on assets (ROA) and market performance data (Tobin's Q). Results show that ESG has a positive impact on return on assets, but find no effect on Tobin's Q. In addition, of the three pillars, governance performance has the strongest effect on financial performance. As explained in the paper, a reason as to why this pillar shows stronger effects, might be because of longer traditions regarding corporate governance reporting. Partly because of regulatory reasons that were introduced in 2002.

4. Hypotheses

There are two main purposes of this study. First, we want to evaluate if the effect of changes in a firm's one-year lagged ESG pillar scores, which are environmental, social and governance, on the firm's stock returns is more pronounced for developed market companies than developing market companies. In other words, do investors require less return for a unit improvement in the ESG score of a developed market company compared to a unit improvement of a developing market company. Second, we want to evaluate if a portfolio of highly rated ESG stocks from developing markets will deliver a risk-adjusted excess return over a portfolio of low rated ESG stocks from developed markets.

The first hypothesis is tested with a panel regression in which annual stock returns are regressed on the one-year lagged ESG pillar scores while controlling for individual firm characteristics. The second hypothesis is tested by creating a portfolio that is long in the top 20% of developing market stocks based on combined ESG score and short in the bottom 20% of developed market stocks based on combined ESG score. The Carhart (1997) four factor regression is run on the monthly portfolio returns to evaluate the portfolio's alpha.

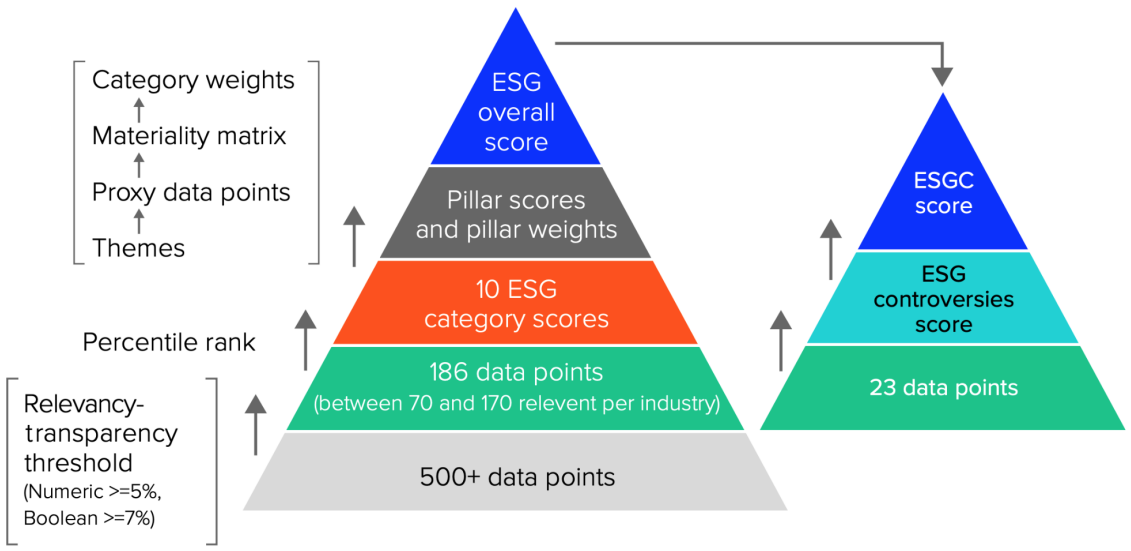
5. Empirical Work

5.1 Data

Data has been collected from multiple sources for our study. The purpose of this section is to give a descriptive overview of this process.

First, a majority of our data is retrieved from an Eikon terminal using the filter or screening application which could be accessed from the terminal itself. The data delivered by the Eikon platform is provided by Refinitiv, a global provider of financial market data and infrastructure (Refinitiv, 2021). Through the screener application, we retrieved the following company specific information. Monthly stock returns are retrieved using "Monthly Total Return". This variable includes price changes and dividends for the specified period. The complete ESG score is collected using the "ESG Combined Score" in short, ESGC. This variable is calculated as the average of the ESG score and ESG controversies score when there are controversies during the fiscal year, according to Refinitiv. The methodology used to create the overall score is illustrated by the picture below (Refinitiv Environmental, Social and Governance Scores, 2021).

Figure 4: The Refinitiv five-step ESG scoring methodology



Data for each separate pillar score is collected using three different variables which are described as follows. Environmental Pillar Score which consists of three categories which are

Emission, Innovation and Resource Use (Appendix 1). This pillar measures the impact a company has on living and non-living natural systems which includes air, land, water and complete ecosystems.

Social Pillar Score includes CSR Strategy, Management and Shareholders. It measures the company's capacity to generate trust and loyalty with its workforce, customers and society at large. This pillar reflects the reputation and health of a company.

Governance Pillar Score includes Community, Human Rights, Product Responsibility and Workforce. This pillar measures the company's systems and processes which in turn assures that its board members and executives act in the best interests of its long term shareholders. All of the categories are weighted and the pillar score represents the sum of the category weighted scores and the weights vary across industries. According to Refinitiv, these pillar weights are normalized to percentages ranging from 0-100 (Appendix 2).

The panel regression also included data from additional factors described in the following way by Refinitiv (Refinitiv Eikon, 2021). *Company Market Cap* is the sum of market value for all the relevant share type issues, including default, free float and outstanding share type. *Total Assets Reported* represents the total assets of a company.

Total Liabilities is the sum of total current liabilities, total long-term debt, deferred income tax, as well as total minority interest and other liabilities. *Operating Income*, the sum of total revenue and total operating expense. *Total Revenue* is the sum of a company's operating activities after deducting sales adjustments and their equivalents. *Dividend Per Share Yield* in percent is the security's dividend per share divided by its price. This measurement represents the company's common stocks dividend on an annual basis which in turn is divided by the weighted average number of common shares outstanding for the year. For US companies the dividend per share is calculated before withholding taxes, however for some non-US companies this measurement is calculated after withholding taxes.

Total Return includes the change in price and any relevant dividends for the specified period. Compounded daily return for the specified period is used to calculate the total return and is based on the dividend reinvested methodology. The dividend type used is the most widely reported for a market and is either gross or net. All of this data was translated to USD for

every company for the purpose of comparability. The previously mentioned data points were used to calculate the control variables for the panel regression.

Second, risk-free rates had to be collected from several sources due to lack of a comprehensive database for the countries included in the portfolios. Collection is mainly done through OECD's website where the indicator name is "Short term interest rate annual %" (OECD 2021). This amount is divided by 12 to obtain the monthly risk-free rate since we used the monthly stock returns in the Carhart (1997) regression. For cases when data was not available for individual countries in the OECD database, we searched for data in the IMF database using the indicator "Government T bills - annual %" (IMF 2021). For few countries that do not issue T-bills, the yield on long term government bonds is used as the benchmark interest rate.

In some instances, when the risk-free data for certain countries and/or periods was not available both in the OECD and IMF datasets, we used the website or databank of the individual country's central bank to obtain the risk-free benchmark data. For each one of these events, the yield for the bill or bond with the shortest duration for which data is available was chosen. This was the case for Taiwan, Kuwait, Oman and Turkey. In some exceptional cases, where we could not find the relevant data from either OECD, IMF, or the country's central bank, we referred to third party data sources like World Government Bonds (World Government Bonds, 2021) or CEIC (CEIC 2021). This was the case for Peru, Jordan and Argentina.

Third, the monthly factor return data for the Carhart (1997) regression is retrieved from Kenneth French's data library. The Carhart (1997) regression uses the three Fama/French factors of market (Rm-Rf), size (SMB) and value (HML) as well as the Momentum (WML) factor. Our study uses both the *developed market factor returns* which consist of stocks from several developed countries[1] and the *emerging market factor returns* which consist of stocks from selected emerging markets[2].

[1] Australia, Austria, Belgium, Canada, Switzerland, Germany, Denmark, Spain, Finland, France, Great Britain, Greece, Hong Kong, Ireland, Italy, Japan, Netherlands, Norway, New Zealand, Portugal, Sweden, Singapore, United States.

[2] Argentina, Brazil, Chile, China, Colombia, Czech Republic, Egypt, Greece, Hungary, India, Indonesia, Malaysia, Mexico, Pakistan, Peru, Philippines, Poland, Qatar, Russia, Saudi Arabia, South Africa, South Korea, Taiwan, Thailand, Turkey, United Arab Emirates.

5.2 Empirical Method

We implemented two distinct empirical methods in order to test the two aforementioned hypotheses. Our first approach applies a panel regression to evaluate whether the effect on stock returns of changes in a firm's one-year lagged ESG pillar scores - environmental, social, and governance - is more pronounced for developed market companies compared to developing market companies. Our second approach applies the Carhart (1997) four factor model to evaluate whether a portfolio that is long in the top 20% ESG-rated stocks from the developing market and short in the bottom 20% ESG-rated stocks from the developed markets will generate a risk-adjusted alpha.

5.2.1 Panel Regression

The panel regression uses the one-year firm-level "total return" from Eikon as the dependent variable and the one-year lagged ESG pillar scores, namely, the environment, social and governance scores as independent variables. In order to evaluate the difference in effects on stock returns of changes in one-year lagged ESG score between developed and developing market companies, we add a *class* dummy variable to the regression. The class variable is 1 for a company from a developed country and 0 otherwise. With regard to control variables, we follow earlier literature and employ the standard firm characteristics for the analyses on the firm level (Capon et al, 1990; Brailsdorf et al, 2002; Titman and Wessels, 1988; Bannier et al, 2019). These characteristics are *Leverage*, *Size*, *Profitability*, *Efficiency* and *Dividend Yield*.

$$R_{i,t+1} - r_{f,t+1} = \beta_1 Class_{t+1} + \beta_2 Envtscore_t + \beta_3 Socscore_{i,t} + \beta_4 Govtscore_{i,t} + \beta_4 X_{i,t} + \epsilon_{i,t}$$

$R_{i,t+1} - r_{f,t+1}$ is the dependent variable and denotes the annual excess return of firm i in year t+1. The stock return for firm i in year t+1 ($R_{i,t+1}$) is mapped to the risk-free rate of firm i's country of headquarters in year t+1 ($r_{f,t+1}$) to obtain the excess returns for firm i ($R_{i,t+1} - r_{f,t+1}$). $Class_{t+1}$ is a dummy variable that is 1 if firm i is from a developed market and 0 otherwise. Taking inspiration from the UN's country classification, we defined the developed market to consist of the EU-15[3] plus Switzerland, Norway, United States, Canada, Japan, Australia and New Zealand. $Envtscore_t$ denotes the one-year lagged environment pillar score for firm i, $Socscore_{i,t}$ denotes the one-year lagged social pillar score for firm i and $Govtscore_{i,t}$ is the

[3] EU-15 include Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden and United Kingdom

one-year lagged governance pillar score for firm i. $X_{i,t}$ is a vector of one-year lagged control variables. Following Bannier et al (2019), the control variable vector consist of firm *Leverage* (proxied as the ratio of total assets to total liabilities), firm *Size* (proxied as the natural logarithm of total sales), firm *Profitability* (proxied by operating income divided by total assets), firm *Efficiency* (proxied by total revenues divided by total assets) and *Dividend Yield*.

The data set for the panel regression consists of 26 745 firm-year observations of companies spread across the 11 industries of the Industry Classification Benchmark (ICB). The number of countries included in the sample range from 12 countries in 2003 to 56 countries in 2020. The subsequent tables provide an overview of the sample distribution across year, country and industry as well as summary statistics of the regression variables.

Table 1: Summary distribution of data across year, country and industry

Year	Companies	Countries	Industries
2003	54	12	11
2004	278	24	11
2005	428	27	11
2006	544	28	11
2007	808	27	11
2008	805	29	11
2009	827	33	11
2010	1162	45	11
2011	1164	44	11
2012	1879	49	11
2013	2004	50	11
2014	1637	47	11
2015	2171	53	11
2016	1882	50	11
2017	2287	53	11
2018	2603	54	11
2019	2955	56	11
2020	3254	56	11

Table 2: Summary statistics of regression variables

	envtscore	socscore	govtscore	divyield	leverage	size	profitability	efficiency	excess_ret
nobs	26745.0	26745.0	26745.0	26745.0	26745.0	26745.0	26745.0	26745.0	26745.0
Maximum	98.9	98.6	99.3	292.5	28335.5	28.6	3.1	7.0	1292.7
Mean	35.5	43.1	49.5	0.0	4.2	22.3	0.1	0.8	10.5
Stdev	29.2	23.7	22.6	1.8	181.3	1.6	0.1	0.6	45.4
Skewness	0.3	0.3	0.0	163.2	144.6	0.1	-2.4	2.0	4.4
Kurtosis	-1.2	-0.8	-0.9	26662.0	22324.8	0.3	162.0	6.8	74.6

5.2.2 Carhart (1997) Four Factor Model

The analysis of difference in stock returns has a long history in financial research (Bannier et al, 2019). Fama and French (1993) identify three main risk factors (market, size and value) that drive stock returns. Carhart (1997) later introduced momentum as a fourth factor. In order to investigate if a portfolio of highly rated ESG stocks from developing markets will deliver a risk-adjusted excess return over a portfolio of low rated ESG stocks from developed markets, we apply the Carhart (1997) four factor regression on a zero-cost portfolio that is long in Q5 (top 20% ESG rating) of developing market stocks and short in Q1 (bottom 20% ESG rating) of developed market stocks. We subsequently run the following regression equation on the monthly returns of the corresponding zero-cost portfolio:

$$R_t - r_t = \alpha + \beta_1 Mkt.RF_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 Mom_t + \epsilon_t$$

$R_t - r_t$ is the dependent variable and denotes the monthly excess returns of the zero-cost portfolios. The excess returns of the portfolio's constituents are used to calculate the portfolio's excess return and each constituent's excess return is computed using the risk-free rate of the specific period in the constituent's country of headquarters. $MKt.RF_t$ is the excess monthly returns on the market portfolio ($R_m - R_f$). SMB_t, HML_t and Mom_t denote the monthly returns of size (small minus big), value (book-to-market) and momentum factors respectively. The regression intercept alpha (α) is our variable of interest, and can be interpreted as the risk-adjusted outperformance of a portfolio of top 20% ESG-rated stocks in the developing markets finance with a short position in the bottom 20% of ESG-rated stocks in the developed markets.

5.2.2.1 Portfolio Construction

To construct the zero-cost portfolio, we classify the universe of stocks between developed and developing countries. Consistent with the approach in section 5.2.1, we defined the

developed market to consist of the EU-15[4] plus Switzerland, Norway, United States, Canada, Japan, Australia and New Zealand. Developing countries are defined as all other countries not classified in the developed country list.

The ESG-combined (ESGC) score on December 30th of year t is used to determine the constituents of the portfolio to be constructed in July of year t+1. The long position consists of the top 20% ESG-rated stocks in developing markets while the short position consists of bottom 20% ESG-rated stocks in developed markets.

Our sample includes both active and inactive companies in order to avoid survivorship bias in the results. The data universe includes 2,889 companies from 53 countries resulting in 117,736 firm-month observations. The developing market accounted for 557 companies, and the developed markets accounted for 2,332 companies.

[4] EU-15 include Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden and United Kingdom.

Table 3: Descriptive Statistics of developed market companies

The table shows summary statistics for the constituents of the short position (bottom 20% ESG-rated developed market stocks). excess_rett is the portfolio's average monthly excess return and sharpe_ratio is the portfolio's average monthly sharpe ratio.

formation_yr	companies	industries	countries	excess_rett	sharpe_ratio
2006	241	11	21	1.72	0.24
2007	289	11	21	-1.71	-0.16
2008	296	11	21	-1.17	-0.06
2009	323	11	21	1.25	0.11
2010	374	11	20	1.75	0.20
2011	407	11	19	-0.27	-0.02
2012	431	11	19	1.81	0.11
2013	445	11	20	1.78	0.17
2014	459	11	18	0.26	0.02
2015	477	11	17	-0.19	-0.01
2016	508	11	18	1.57	0.14
2017	652	11	18	1.16	0.11
2018	796	11	17	-0.30	-0.02
2019	884	11	19	0.28	0.01
2020	1002	11	21	5.25	0.25

Table 4: Descriptive statistics of developing market companies

The table shows summary statistics for the constituents of the long position (top 20% ESG-rated developing market stocks). excess_rett is the portfolio's average monthly excess return and sharpe_ratio is the portfolio's average monthly sharpe ratio.

formation_yr	companies	industries	countries	excess_rett	sharpe_ratio
2006	13	6	5	1.92	0.32
2007	15	7	5	0.42	0.04
2008	19	9	9	0.04	0.00
2009	30	10	11	1.46	0.17
2010	68	10	19	1.48	0.21
2011	105	11	20	-0.37	-0.04
2012	171	11	22	0.24	0.03
2013	204	11	21	0.84	0.11
2014	215	11	23	-0.42	-0.05
2015	227	11	23	-0.14	-0.01
2016	242	11	24	1.13	0.14
2017	262	11	26	0.15	0.02
2018	283	11	30	0.24	0.03
2019	349	11	29	-0.77	-0.06
2020	392	11	32	3.04	0.24

Due to the small sample size from the developing market in the initial years, we started constructing the zero-cost portfolio from 2010 where we had 68 companies from developing markets. For each year from 2010, an equal-weighted (EW) and value-weighted (VW) portfolio of developing and developed markets was constructed on July 1st using constituents determined by the ESG combined score in the preceding December 31st. Each portfolio is held for a period of one (1) year resulting in the so-called 12-6-12 portfolio construction strategy. The one year holding period is chosen taking into account the long-term nature of the effects of sustainability as well as the fact that the ESG scores are published on an annual basis.

Monthly returns of the short position are subtracted from the monthly returns of the long position to obtain the monthly returns of the zero-cost portfolio used in the Carhart (1997) regression. We constructed two zero-cost portfolios long in top 20% ESG-rated stocks from developing markets and short in the bottom 20% ESG-rated stocks from developing markets. The first portfolio is an Equal-weighted portfolio (EWport_ret) and the second is a Value-weighted portfolio (VWport_ret). See appendix for plots of the monthly returns of the long, short and zero-cost portfolios.

5.3 Results

5.3.1 Panel Regression Results

We run three panel regressions to evaluate whether the effect on stock returns of changes in a firm's one-year lagged ESG pillar scores - environmental, social, and governance - is more pronounced for developed market companies compared to developing market companies.

Regression (1) applies a simple pooled OLS method and shows a positive coefficient of 2.5% for the *class* dummy variable, significant at 0.01 level. This indicates that controlling for ESG pillar scores and firm characteristics, stocks in developed markets on average have 2.5% higher return than those in the developing markets. The coefficient of the envtscore score of -0.03 significant at 5% indicates that an increase in envt pillar score results in a decrease in the firm's return in the subsequent year consistent with risk-return tradeoff. The socscore coeff of 0.05 significant at 1% however indicates a positive correlation between social pillar and firm's return. The govt pillar score of -0.01 is not significant.

Regression (2) applies a pooled OLS with time fixed effects and shows a larger coefficient of 3.8% for the *class* dummy variable which is significant at 0.01 level. Controlling for time fixed effects increases the impact of changes in ESG pillar scores on the firm's subsequent year return.

Regression (3) applies a pooled OLS with time and industry fixed effects and shows a positive coefficient of 3.5% for the *class* dummy variable significant at 0.01 level. The panel regressions therefore do not provide evidence to support our first hypothesis. Table 5 summarises the results from the three panel regressions. Controlling for both time and industry specific effects diminished the impact of ESG pillar scores on returns. Furthermore, only the socscore is significant in this case.

Table 5: Output of the three panel regressions

The table presents the output of the three panel regressions. The independent variables are class (1=developed, 0=developing), one-year lagged of environment, social, governance score, leverage, size, profitability, efficiency, and dividend yield. The dependent variable is annual stock return. Regression (1) is a simple pooled OLS, Regression (2) controls for time fixed effect and Regression (3) controls for time and industry fixed effects. Standard errors reported in the parentheses. ***p<0.01, **p<0.05, * p<0.1.

	Dependent Variable:		
	Excess return		
	OLS (1)	Timfix (2)	Tim+ind fix (3)
class	2.480***	3.808***	3.453***
	(0.686)	(0.656)	(0.658)
envtscore	-0.029**	-0.043***	-0.020
	(0.014)	(0.014)	(0.014)
socscore	0.051***	0.065***	0.040**
	(0.017)	(0.016)	(0.017)
govtscore	-0.006	-0.006	-0.006
	(0.014)	(0.013)	(0.013)
leverage	-0.0001	-0.0001	0.00000
	(0.002)	(0.001)	(0.001)
size	-0.807***	-0.932***	-0.808***
	(0.204)	(0.197)	(0.206)
profitability	-7.136***	-6.747***	-7.569***
	(2.224)	(2.093)	(2.097)
efficiency	2.590***	2.432***	2.415***
	(0.462)	(0.434)	(0.486)
divyield	0.207	0.156	0.155
	(0.155)	(0.145)	(0.145)
constant	24.258***		
	(4.480)		
Observations	26,745	26,745	26,745
Adjusted R2	0.003	0.127	0.133
Note:			*p<0.1 **p<0.05 ***p<0.01

We apply the Breusch and Pagan (1979) test on all three regressions to evaluate whether the regression residuals are subject to heteroskedasticity. The BP test on all three regressions returns a p-value < 1% and thus the null hypothesis of homoscedasticity is rejected. We then re-estimate all three regressions with Robust Standard Errors. The results from all three regressions remain unchanged after re-estimating with Huber-White robust standard errors. See appendix 3 for the results with robust standard errors.

5.3.2 Carhart (1997) Four Factor Regression

We run four Carhart (1997) regressions to investigate if a portfolio of highly rated ESG stocks from developing markets will deliver a risk-adjusted excess return over a portfolio of low rated ESG stocks from developed markets.

Regression (1) has the equal-weighted monthly portfolio returns as the dependent variable and monthly *developed market factor* returns as independent variables. Regression (2) has the equal-weighted monthly portfolio returns as the dependent variable and monthly *emerging market factor* returns as independent variables. Regression (1) shows a negative alpha of -5% p.a. significant at 0.1 level, and Regression (2) shows a negative alpha of -9.9% p.a. significant at 0.01 level.

Regression (3) has the value-weighted monthly portfolio returns as the dependent variable and monthly *developed market factor* returns as independent variables. Regression (4) has the value-weighted monthly portfolio returns as the dependent variable and monthly *emerging market factor* returns as independent variables. Regression (3) shows a negative alpha of -5% p.a. significant at 0.1, and Regression (4) shows a negative alpha of -10.9% p.a. significant at 0.01. The alphas p.a. are calculated from the monthly alphas in the regression output.

All four regressions indicate that a portfolio that is long in top 20% ESG-rated stocks from developing markets significantly underperforms a portfolio consisting of the bottom 20% ESG-rated stocks from developed markets. We therefore do not find evidence to support our second hypothesis. The table summarises the results from the four regressions.

Table 6: Output of Carhart (1997) four factor Regression

The table presents the four-factor regressions of equally and value-weighted monthly returns for a portfolio long in top 20% ESG-rated developing market stocks and short in the bottom 20% ESG-rated developed market stocks. The independent variables are excess return on the market (Mkt.RF), size (SMB), value (HML) and momentum (WML). Regression (1) and (4) uses the developed market four-factor returns while Regression (2) and (3) uses emerging market four-factor returns. Standard errors in parentheses. ***p<0.01, **p<0.05, * p<0.1

	Dependent Variable:			
	EW DEV (1)	EW EM (2)	VW DEV (3)	VW EM (4)
Mkt.RF	-0.247***	-0.1**	-0.315***	-0.084
	(0.055)	(0.049)	(0.059)	(0.055)
SMB	-0.490***	-0.652***	-0.127	-0.591***
	(0.153)	(0.152)	(0.163)	(0.170)
HML	0.018	0.478***	0.336**	0.396**
	(0.122)	(0.147)	(0.130)	(0.164)
WML	-0.033	0.182	0.019	0.118
	(0.101)	(0.116)	(0.107)	(0.129)
Constant	-0.426*	-0.865***	-0.424*	-0.961***
	(0.226)	(0.254)	(0.240)	(0.283)
Observations	128	128	128	128
Adjusted R2	0.211	0.157	0.229	0.096
Note:		*p<0.1	**p<0.05	***p<0.01

6. Conclusion

The purpose of this section is to summarize our findings and the impact for investors as well as discuss these findings in relation to other previous studies. In addition, we elaborate on the limitations of our study and the contribution to academia and finally the possibilities for future research.

6.1 Main findings

The main finding of this study is that developed market stocks have a higher risk-adjusted return over developing stocks when ESG risk is controlled for. Furthermore, a portfolio long in high ESG-rated developing market stocks and short in low ESG-rated developed market stocks return a significant negative alpha. These findings do not support the two hypotheses described in section 4.

6.2 Discussion

6.2.1 Implications for investors

The results presented in the previous section highlights several important aspects that align with previous literature and the behavior of investors, both institutional and retail.

From a behavioral aspect, even though returns for highly rated ESG companies might not be superior over those qualified as low ESG as shown in our results. It does not necessarily mean that these types of investments would not attract investors. Since the investor universe consists of a part classified as Type-A investors, these actors will weigh the return and ESG rating in a completely different way than those classified as Type-U (Petersen et al, 2020). Since the aware investors are willing to sacrifice returns in order to do good (Hong and Kacperczyk, 2009), this type of person would still be intrigued to invest in highly rated ESG companies in developing markets. A behavior which would not align with that of the unaware investor who would instead choose the low rated ESG stocks in the developed market to achieve the highest return. However, since the time series is still somewhat limited and the ESG framework relatively young, the results of this study are subject to change when time series expand and the framework becomes more comprehensive.

6.2.2 Comparison to other findings in literature

Other findings in literature suggest similar results to those established in our study. A highly concentrated portfolio of ESG stocks does not generate excess returns (Halbritter and Dorfleitner, 2014). Vlassis (2020) on the other hand presents results that are opposite from those of our study. The main difference in these studies is related to the data used. Vlassis (2020) uses companies in the S&P 500 while Halbritter and Dorfleitner (2014) use data from 8500 companies worldwide in multiple countries. Thus, making direct comparisons to these works is difficult due to the differences in the dataset as well as differences to our study in terms of restrictions set to identify developing and developed countries.

Furthermore, the results align with the statements made in the paper written by Fatemi et al (2018). Since the companies chosen in developing markets are highly rated in terms of ESG, there is a risk that these companies underperform those of low rated ESG companies in developed markets due the market's prediction of overinvestment in sustainability. Due to this overinvestment, companies are "punished" by the market and thus obtain a lower valuation since these overinvestments are predicted to have a neutral or negative effect on the company in the future.

Consistent with the findings of our study, Bodhanwala and Bodhanwala (2019) also finds that SRI portfolios from developing nations significantly underperformed their benchmark index but also carried significantly lower risk compared to the benchmark. The study, however, covers only 64 companies from 5 developing countries and used country-level benchmarks, making it hard to draw a comparison with our study.

6.2.3 Limitations

When conducting our study, there were several limitations that possibly had an effect on our work and in turn the outcome of the project. First, our timespan includes returns from 2006 to 2020 which includes periods of large drawdowns in stock market indexes and equity prices such as the financial crisis in 2008 and the 2020 crash initiated by COVID-19. It is possible that these extreme values had an effect on our regressions and in turn the results that we present in our study.

Second, the availability of ESG rating data is slim in terms of included companies as well as for the available timespan. Many of the rating companies started collecting this type of data around 2000-2005 which explains the somewhat narrow timeframe. Also, the collection of ESG data has been much more extensive in developed countries, thus creating a comprehensive dataset, while the data for developing countries is much narrower. Therefore, we have had to limit our timespan to make sure that the amount of companies included, both from developing and developed markets is sufficient to create a representative portfolio.

Third, due to availability we chose to use the ASSET4 ESG score accessed through an Eikon terminal. As mentioned previously, there are several suppliers of similar ESG scoring systems which might have different datasets in terms of included companies and timespan. An optimal scenario would be to combine all of the scoring systems, creating a more comprehensive dataset. However, this would not have been possible for the scope of this study, due to time constraints as well as limited accessibility to terminals such as Bloomberg. Also, since every scoring system is different, it would require systematic translation to a single system which in turn could yield faulty scores if not done correctly.

Lastly, all of the ESG scoring systems are different in the way that they are created. Both in terms of weighting and inclusion of different factors that in turn make up the total score given to a company. Since there is no universal standard for the ESG measurement, it makes comparability difficult and reliability decreases. Since the metrics are inconsistent in terms of weighting, selection and definition it is difficult to state which of these measurements best captures ESG comparable companies. Also, since there are big differences in collected data in developing and developed countries, there are big differences in the available dataset which in turn could affect the results in our study and lead to skewed results.

6.3 Contribution to academia

Compared to previous studies on the effect of ESG rating on risk and return of developing market stocks, to the best of our knowledge, our study provides the most representative insight on the relationship between ESG-rating and firm stock performance in developing markets than any previous research. Our zero-cost portfolio included 557 companies from 32 developing countries while our panel regression included 1512 companies from 39 developing countries. On the other hand, Bodhanwala and Bodhanwala (2019) which assess the stock market performance and risk of highly compliant firms portfolio in developed and

developing markets included only Brazil, Russia, India and China (BRIC) as developing countries and only 64 companies. Hill et al (2007) which studies SRI portfolio performance in global context for the period 1995 to 2005 included only 10 companies from 4 Asian countries as the developing market.

Unlike most studies that only use *equal-weighted* portfolios (Hill et al, 2007; Bodhanwala and Bodhanwala, 2019; Bannier et al, 2019) to assess jensen alpha of ESG-rated portfolios, we constructed both *equal-weighted* and *value-weighted* portfolios to investigate whether going long in top ESG-rated stocks and financing it with a short position in bottom ESG-rated stocks will generate an alpha. Our results are consistent across both portfolio strategies.

6.4 Future research

After conducting our study, some of our conclusions have led to a number of additional questions that are relevant for future research.

First, given access to a broader timespan of data, mainly for both the ESG score and the individual ESG pillar scores, the same study could be performed. Expanding the timeframe could potentially yield different results compared to those received in this study. Also, a more selective process for certain time periods such as removing years when crashes occurred would also be suitable for further research.

Second, performing a similar study by combining multiple ESG scores would also be relevant to study further. In this study an ESG score from Refinitiv was used, however, combining multiple sources of ESG measurements from sources such as Bloomberg and Fitch would provide an additional dimension in the form of a more comprehensive dataset.

Lastly, elaborating with selection for factors such as countries, quintiles and control variables might also provide an interesting topic for future studies. For example, choosing another limit for the quintiles, either a more selective restriction (less than 20%) or a broder restriction (more than 20%).

7. References

7.1 Articles

Berry, T. C., & Junkus, J. C. (2013). Socially responsible investing: An investor perspective. *Journal of business ethics*, *112*(4), 707-720.

Dorfleitner, G., Halbritter, G., & Nguyen, M. (2015). Measuring the level and risk of corporate responsibility–An empirical comparison of different ESG rating approaches. *Journal of Asset Management*, *16*(7), 450-466.

Escrig-Olmedo, E., Fernández-Izquierdo, M. Á., Ferrero-Ferrero, I., Rivera-Lirio, J. M., & Muñoz-Torres, M. J. (2019). Rating the raters: Evaluating how ESG rating agencies integrate sustainability principles. *Sustainability*, *11*(3), 915.

Vlassis, N. (2020). ESG Investing: In Pursuit of Alpha.

Drempetic, S., Klein, C., & Zwergel, B. (2019). The influence of firm size on the ESG score: Corporate sustainability ratings under review. *Journal of Business Ethics*, 1-28.

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

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

Bannier, C. E., Bofinger, Y., & Rock, B. (2019). *Doing safe by doing good: ESG investing and corporate social responsibility in the US and Europe* (No. 621). CFS Working Paper Series.

Capon, N., J. U. Farley, and S. Hoenig (1990): *Determinants of Financial Performance: A Meta-Analysis*. Management Science, 36(10), 1143–1159

Brailsford, T. J., B. R. Oliver, and S. L. H. Pua (2002): *On the relation between ownership structure and capital structure*. Accounting and Finance, 42(1), 1–26

Titman, S., and R. Wessels (1988): *The Determinants of Capital Structure Choice*. The Journal of Finance, 43(1), 1–19

Riedl, A., & Smeets, P. (2017). Why do investors hold socially responsible mutual funds?. *The Journal of Finance*, *72*(6), 2505-2550.

Hong, H., & Kacperczyk, M. (2009). The price of sin: The effects of social norms on markets. *Journal of financial economics*, *93*(1), 15-36.

Anderson, A., & Robinson, D. T. (2019). Talking About the Weather: Availability, Affect and the Demand for Green Investments. *Swedish House of Finance Research Paper*, (19-14).

Pedersen, L. H., Fitzgibbons, S., & Pomorski, L. (2020). Responsible investing: The ESG-efficient frontier. Journal of Financial Economics.

Odell, J., & Ali, U. (2016). ESG investing in emerging and frontier markets. *Journal of Applied Corporate Finance*, *28*(2), 96-101.

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*, *150*, 135-147.

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

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

Hartzmark, S. M., & Sussman, A. B. (2019). Do investors value sustainability? A natural experiment examining ranking and fund flows. *The Journal of Finance*, *74*(6), 2789-2837.

Margolis, J. D., Elfenbein, H. A., & Walsh, J. P. (2009). Does it pay to be good... and does it matter? A meta-analysis of the relationship between corporate social and financial performance. And does it matter.

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

Hill, R. P., Ainscough, T., Shank, T., & Manullang, D. (2007). Corporate social responsibility and socially responsible investing: A global perspective. *Journal of business ethics*, *70*(2), 165-174.

Breusch, T.S., and A.R. Pagan. 1979. *A simple test for heterosce- dasticity and random coefficient variation*. Econometrica 47: 1287–1294

7.2 Links and websites

Morningstar, (2020), *Global Sustainable Fund Flows: Q4 2020 in Review*, viewed May 1 2021
https://www.morningstar.com/content/dam/marketing/shared/pdfs/Research/Global_ESG_Q4_2020_Flows.pdf?utm_source=eloqua&utm_medium=email&utm_campaign=&utm_content=27223

EUROSIF, (2018), *European SRI study 2018*, viewed May 1 2021
http://www.eurosif.org/wp-content/uploads/2018/11/European-SRI-2018-Study.pdf

Refinitiv 2021, *Refinitiv ESG company scores methodology,* viewed April 1 2021
https://www.refinitiv.com/en/sustainable-finance/esg-scores#methodology

Refinitiv 2021, *About Us*, viewed April 1 2021
https://www.refinitiv.com/en/about-us

Refinitiv, (2021), *Environmental, Social and Governance (ESG) Scores*, viewed April 1 2021
https://www.refinitiv.com/content/dam/marketing/en_us/documents/methodology/refinitiv-esg-scores-methodology.pdf,

World Government Bonds, 2021
http://www.worldgovernmentbonds.com/

IMF, 2021
https://data.imf.org/regular.aspx?key=61545867

Organisation for Economic Co-operation and Development (OECD), *Long-term Interest Rates* 2021
https://data.oecd.org/interest/short-term-interest-rates.htm

Kenneth R French Website
http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/index.html

8. Appendix

8.1 Refinitiv ESG Materiality Matrix

Pillars	Catagories	Themes	Data points	Weight method
Environmental	Emmission	Emissions	TR.AnalyticCO2	Quant industry median
		Waste	TR.AnalyticTotalWaste	Quant industry median
		Biodiversity *		
		Environmental management systems *		
	Innovation	Product innovation	TR.EnvProducts	Transparency weights
		Green revenues/R&D/capex	TR.AnalyticEnvRD	Quant industry median
	Resource use	Water	TR.AnalyticWaterUse	Quant industry median
		Energy	TR.AnalyticEnergyUse	Quant industry median
		Sustainable packaging *		
		Environmental supply chain *		
Governance	CSR strategy	CSR strategy	Data points in governance category/data points in governance pillar	Count of data points in each governance category/All data points in governance pillar
		ESG reporting and transparency		
	Management	Structure (independence, diversity, committees)	Data points in governance category/data points in governance pillar	Count of data points in each governance category/All data points in governance pillar
		Compensation		
	Shareholders	Shareholder rights	Data points in governance category/data points in governance pillar	Count of data points in each governance category/All data points in governance pillar
		Takeover defenses		
Social	Community	Equally important to all industry groups, hence a median weight of 5 is assigned to all industry groups		Equally important to all industry groups
	Human rights	Human rights	TR.PolicyHumanRights	Transparency weights
	Product responsibility	Responsible marketing	TR.PolicyResponsibleMarketing	Transparency weights
		Product quality	TR.ProductQualityMonitoring	Transparency weights
		Data privacy	TR.PolicyDataPrivacy	Transparency weights
	Workforce	Diversity and inclusion	TR.WomenEmployees	Quant industry median
		Career development and training	TR.AvgTrainingHours	Transparency weights
		Working conditions	TR.TradeUnionRep	Quant industry median
		Health and safety	TR.AnalyticLostDays	Transparency weights

* No data points available that may be used as a proxy for ESG magnitude/materiality

8.2 Pillar Scoring Matrix

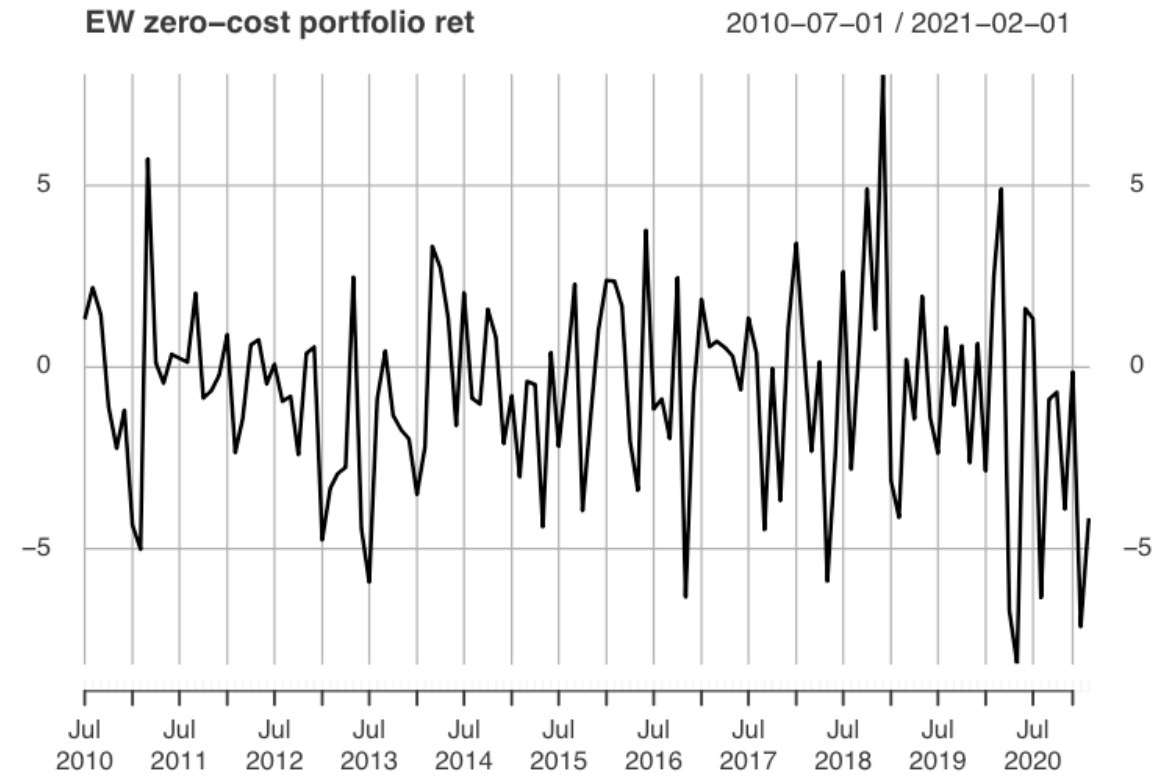
	Environmental			Social				Governance		
Industry group	Emission	Innovation	Resource use	Human rights	Product responsibility	Workforce	Community	Management	Shareholders	CSR strategy
Aerospace and defense	0.36	0.36	0.27	0.34	0.16	0.25	0.25	0.67	0.20	0.13
Automobiles and auto parts	0.29	0.48	0.24	0.35	0.21	0.24	0.20	0.67	0.20	0.13
Banking services	0.17	0.67	0.17	0.19	0.18	0.39	0.24	0.67	0.20	0.13
Beverages	0.42	0.14	0.44	0.33	0.27	0.23	0.18	0.67	0.20	0.13
Biotechnology and medical research	0.35	0.10	0.55	0.07	0.34	0.22	0.37	0.67	0.20	0.13
Chemicals	0.32	0.34	0.34	0.39	0.18	0.23	0.20	0.67	0.20	0.13
Coal	0.49	0.05	0.46	0.22	0.07	0.35	0.36	0.67	0.20	0.13
Collective investments	0.33	0.33	0.33	0.09	0.21	0.25	0.45	0.67	0.20	0.13
Communications and networking	0.24	0.47	0.29	0.13	0.36	0.17	0.34	0.67	0.20	0.13
Computers, phones and household electronics	0.24	0.60	0.16	0.37	0.22	0.21	0.19	0.67	0.20	0.13
Construction and engineering	0.38	0.38	0.24	0.34	0.13	0.29	0.24	0.67	0.20	0.13
Construction materials	0.36	0.28	0.36	0.33	0.12	0.32	0.23	0.67	0.20	0.13
Containers and packaging	0.37	0.24	0.39	0.39	0.18	0.23	0.20	0.67	0.20	0.13
Diversified industrial goods wholesalers	0.21	0.50	0.29	0.43	0.05	0.29	0.24	0.67	0.20	0.13
Diversified retail	0.42	0.12	0.46	0.13	0.28	0.26	0.33	0.67	0.20	0.13
Electric utilities and ipps	0.37	0.29	0.33	0.21	0.15	0.38	0.26	0.67	0.20	0.13
Electronic equipment and parts	0.33	0.28	0.38	0.31	0.12	0.26	0.31	0.67	0.20	0.13
Food and drug retailing	0.46	0.21	0.33	0.21	0.32	0.26	0.21	0.67	0.20	0.13
Food and tobacco	0.45	0.09	0.45	0.27	0.30	0.24	0.19	0.67	0.20	0.13
Freight and logistics services	0.36	0.33	0.31	0.21	0.21	0.33	0.26	0.67	0.20	0.13
Healthcare equipment and supplies	0.36	0.36	0.29	0.25	0.30	0.19	0.25	0.67	0.20	0.13
Healthcare providers and services	0.40	0.13	0.47	0.17	0.33	0.22	0.28	0.67	0.20	0.13
Homebuilding and construction supplies	0.28	0.44	0.28	0.34	0.21	0.23	0.21	0.67	0.20	0.13
Hotels and entertainment services	0.44	0.07	0.48	0.18	0.40	0.20	0.22	0.67	0.20	0.13
Household goods	0.30	0.48	0.23	0.31	0.24	0.23	0.22	0.67	0.20	0.13
Industrial conglomerates	0.30	0.41	0.30	0.37	0.19	0.24	0.20	0.67	0.20	0.13
Insurance	0.20	0.60	0.20	0.17	0.23	0.31	0.29	0.67	0.20	0.13
Investment banking and investment services	0.22	0.56	0.22	0.08	0.21	0.32	0.39	0.67	0.20	0.13
Investment holding companies	0.41	0.09	0.50	0.20	0.10	0.22	0.49	0.67	0.20	0.13
Leisure products	0.27	0.45	0.27	0.19	0.43	0.13	0.24	0.67	0.20	0.13
Machinery, tools, heavy vehicles, trains and ships	0.25	0.53	0.22	0.31	0.24	0.20	0.25	0.67	0.20	0.13
Media and publishing	0.40	0.30	0.30	0.20	0.30	0.24	0.25	0.67	0.20	0.13
Metals and mining	0.44	0.09	0.47	0.41	0.09	0.29	0.20	0.67	0.20	0.13
Multiline utilities	0.35	0.31	0.33	0.28	0.19	0.30	0.23	0.67	0.20	0.13
Natural gas utilities	0.33	0.33	0.35	0.22	0.22	0.35	0.22	0.67	0.20	0.13
Office equipment	0.22	0.59	0.19	0.37	0.26	0.19	0.18	0.67	0.20	0.13
Oil and gas	0.32	0.30	0.39	0.37	0.15	0.29	0.19	0.67	0.20	0.13
Oil and gas related equipment and services	0.46	0.14	0.40	0.38	0.11	0.28	0.23	0.67	0.20	0.13
Paper and forest products	0.33	0.32	0.35	0.32	0.05	0.37	0.26	0.67	0.20	0.13
Passenger transportation services	0.39	0.18	0.42	0.23	0.20	0.35	0.23	0.67	0.20	0.13
Personal and household products and services	0.37	0.27	0.37	0.27	0.33	0.22	0.17	0.67	0.20	0.13
Pharmaceuticals	0.41	0.14	0.45	0.31	0.21	0.26	0.22	0.67	0.20	0.13
Professional and commercial services	0.33	0.29	0.38	0.31	0.19	0.23	0.26	0.67	0.20	0.13
Real estate operations	0.38	0.24	0.38	0.12	0.12	0.45	0.30	0.67	0.20	0.13
Renewable energy	0.35	0.32	0.32	0.08	0.25	0.25	0.42	0.67	0.20	0.13
Residential and commercial REITs	0.43	0.11	0.46	0.07	0.24	0.34	0.36	0.67	0.20	0.13
Semiconductors and semiconductor equipment	0.31	0.39	0.31	0.36	0.20	0.23	0.20	0.67	0.20	0.13
Software and IT services	0.22	0.44	0.33	0.15	0.28	0.17	0.39	0.67	0.20	0.13
Specialty retailers	0.40	0.27	0.33	0.18	0.28	0.24	0.30	0.67	0.20	0.13
Telecommunications services	0.35	0.30	0.35	0.27	0.31	0.26	0.17	0.67	0.20	0.13
Textiles and apparel	0.26	0.37	0.37	0.29	0.25	0.28	0.18	0.67	0.20	0.13
Transport infrastructure	0.42	0.13	0.45	0.26	0.13	0.38	0.22	0.67	0.20	0.13
Uranium	0.48	0.05	0.48	0.08	0.26	0.26	0.40	0.67	0.20	0.13
Water and related utilities	0.35	0.29	0.35	0.17	0.13	0.43	0.28	0.67	0.20	0.13

8.3 Panel regression with robust SE

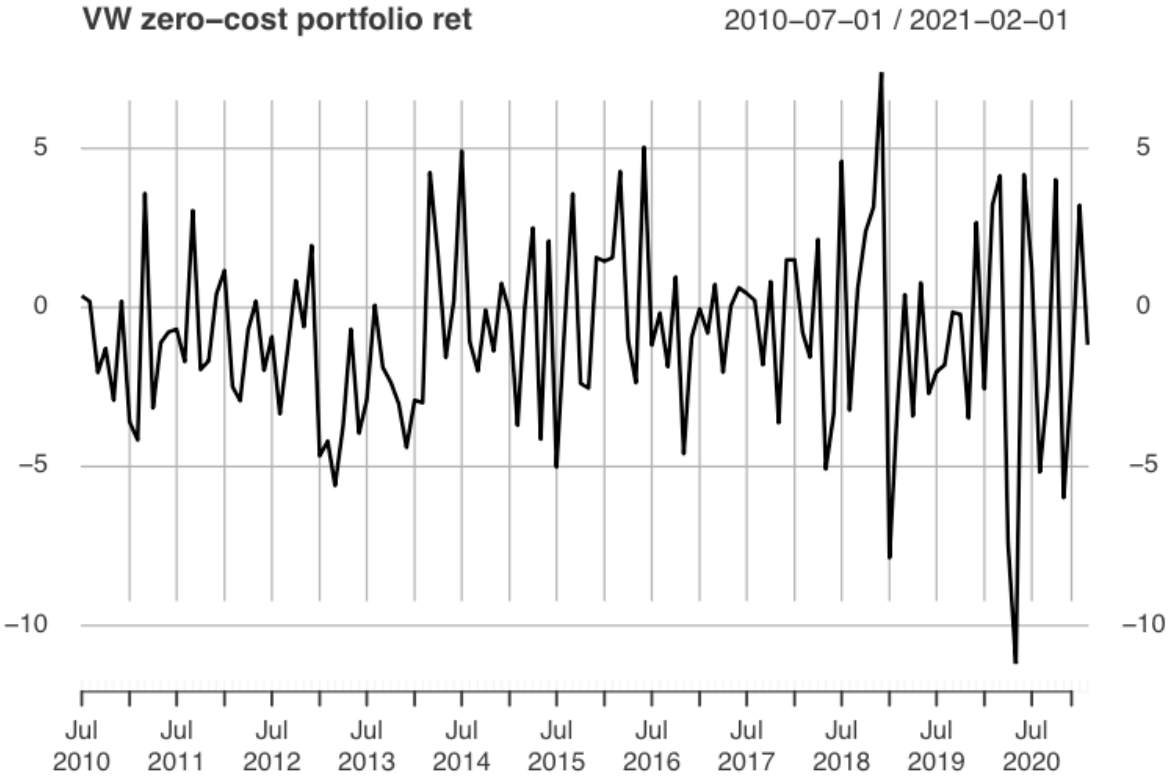
The table presents the output of the three panel regressions estimated with robust standard errors (in parentheses). The independent variables are class (1=developed, 0=developing), one-year lagged of environment, social, governance score, leverage, size, profitability, efficiency, and dividend yield. The dependent variable is annual stock return. Regression (1) is a simple pooled OLS, Regression (2) controls for time fixed effect and Regression (3) controls for time and industry fixed effects. Robust Standard Errors reported in the parentheses. ***p<0.01, **p<0.05, * p<0.1.

	Dependent variable:		
	OLS RSE	timfix RSE	tim+Ind RSE
	(1)	(2)	(3)
class	2.480***	3.808***	3.453***
	(0.685)	(0.670)	(0.674)
envtscore	-0.029**	-0.043***	-0.020
	(0.014)	(0.013)	(0.014)
socscore	0.051***	0.065***	0.040**
	(0.019)	(0.018)	(0.018)
govtscore	-0.006	-0.006	-0.006
	(0.014)	(0.013)	(0.013)
leverage	-0.0001	-0.0001	0.00000
	(0.0004)	(0.0004)	(0.0005)
size	-0.807***	-0.932***	-0.808***
	(0.232)	(0.224)	(0.229)
profitability	-7.136*	-6.747*	-7.569**
	(3.712)	(3.628)	(3.606)
efficiency	2.590***	2.432***	2.415***
	(0.547)	(0.518)	(0.573)
divyield	0.207***	0.156***	0.155***
	(0.023)	(0.011)	(0.015)
Constant	24.258***		
	(5.223)		
Note:	*p<0.1; **p<0.05; ***p<0.01		

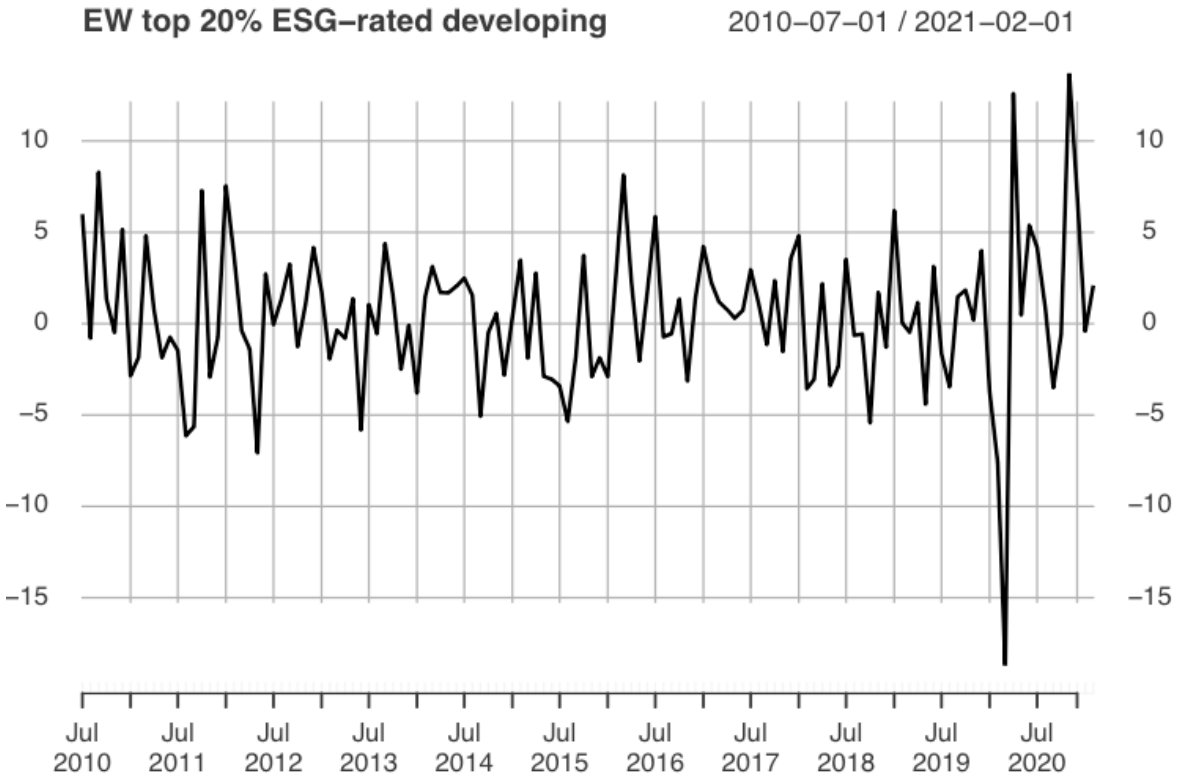
8.4 *Equal-weighted* zero-cost portfolio return (%)



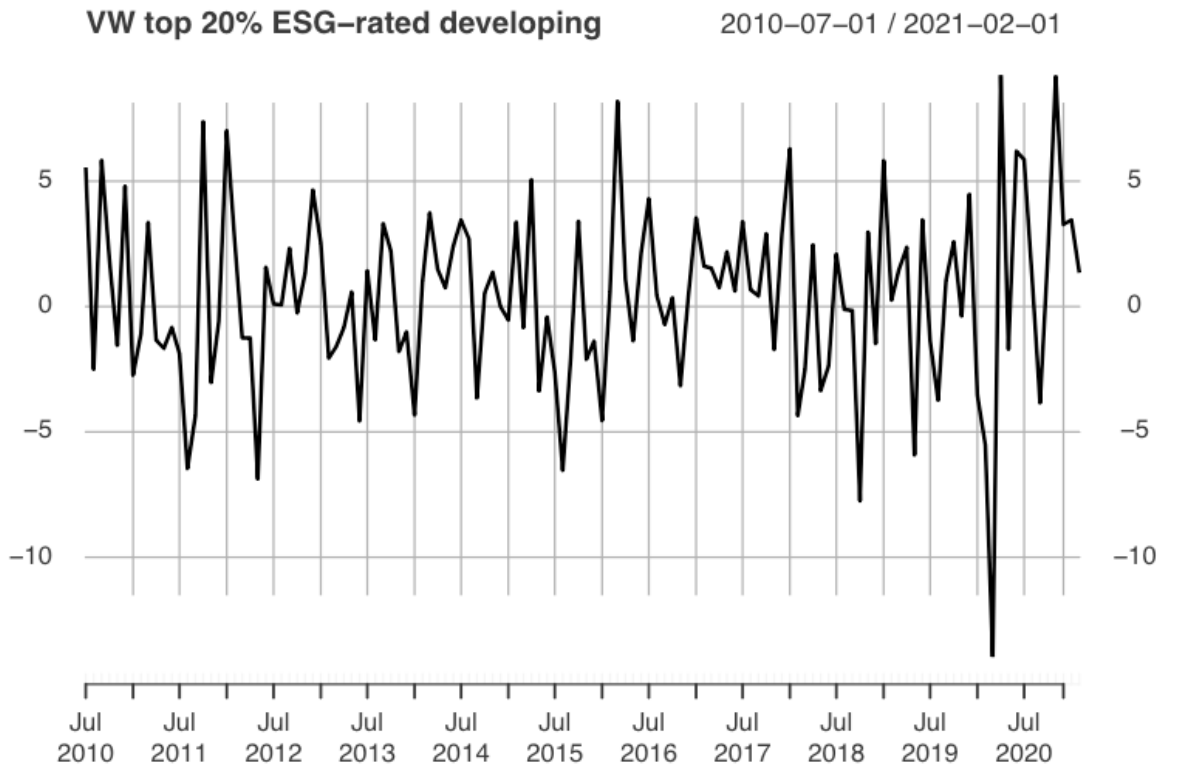
8.5 *Value-weighted* zero-cost portfolio return (%)



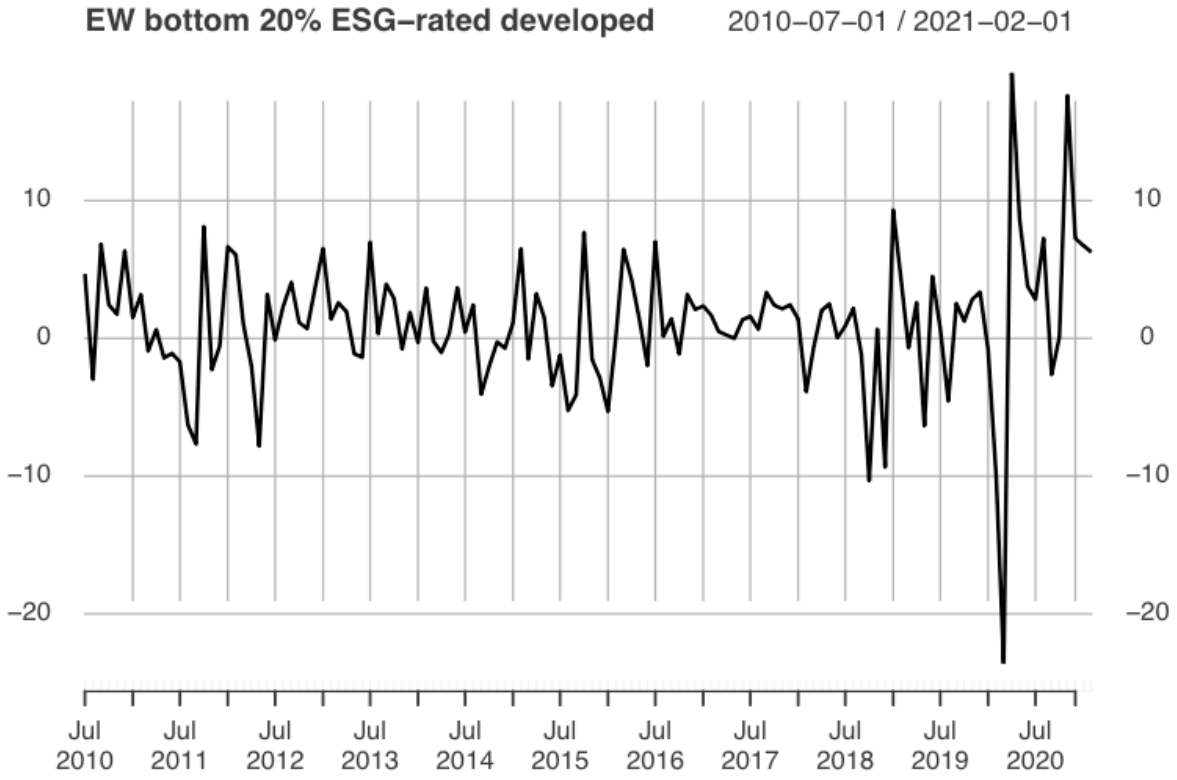
8.6 *Equal-weighted* long portfolio return (%)



8.7 *Value-weighted* long portfolio return (%)



8.8 *Equal-weighted* short portfolio return (%)



8.9 *Value-weighted* short portfolio return (%)

