Norwegian School of Economics
Bergen, Fall 2022

ESG Ratings and Financial Performance

An Empirical Investigation of the Link Between ESG Ratings and Stock Performance in Emerging Markets

Fredrik Thorsen Kulsrud and Nicolai Storvik

Supervisor: Petter Bjerksund

Master thesis, Economics and Business Administration

Major: Financial Economics and Business Analysis and Performance Management

NORWEGIAN SCHOOL OF ECONOMICS

Acknowledgements

This thesis was written as the final part of our Master of Science in Economics and Business Administration at the Norwegian School of Economics, majoring in Financial Economics and Business Analysis and Performance Management.

The importance of sustainability in finance has been particularly pronounced in developed markets, while emerging markets have traditionally lagged behind. Consequently, the large body of existing literature has predominantly been centered around developed markets, leaving emerging markets an underresearched field. This void sparked our curiosity on the subject matter, and through our thesis, we aim to expand on the current understanding of sustainable investing in emerging markets.

While writing our thesis, we have acquired a deeper insight into the role of responsible investments and sustainability in finance. Moreover, our thesis has required extensive knowledge of financial theory and econometric methods, deep insight into Python data science techniques, and knowledge of Microsoft Excel and LaTeX.

We want to extend our sincere gratitude to our supervisor Petter Bjerksund for his valuable and constructive feedback during the development and writing of our thesis. In addition, we want to emphasize our appreciation for his flexible availability and swift responses throughout this process. Lastly, we want to express our gratitude to the Norwegian School of Economics for providing the necessary access to the data sources used for our thesis.

Norwegian School of Economics

Bergen, December 2022

Fredrik Thorsen Kulsrud

Nicolai Storvik

Abstract

Our study aimed to investigate the relationship between Environmental, Social, and Governance (ESG) performance and financial performance in emerging markets. We collected ESG scores from three different rating providers (Bloomberg, Refinitiv, and Sustainalytics), which were used to construct two portfolios consisting of high- and low ESG-rated firms, and a resulting high *minus* low (high-low difference) portfolio. We hypothesized that investors in emerging markets would pay a premium for holding high ESG-rated firms over low ESG-rated firms. To test this hypothesis, we applied the Fama-French framework to examine historical stock returns for each portfolio over the four years from 2018 to 2021. While we observed statistically and economically significant results for both the high- and low portfolio, we did not find any significant monthly abnormal returns for the high-low difference portfolio. This suggests that we are unable to conclusively reject our hypothesis about the relationship between ESG performance and financial performance in emerging markets. However, our study still contributes to a better understanding of this underresearched area and provides valuable insights into the performance of high- and low ESG-rated firms during a period of high market volatility.

Keywords – Sustainability, ESG, Stock Performance, Emerging Markets

Contents

List of Figures

List of Tables

1 Introduction

Traditionally, analysis of corporate financial performance has predominantly revolved around the management of company resources, methods of improving profitability, and overall shareholder value. However, in the current investment climate, firm sustainability performance has, to a greater extent, emerged to be considered a supplementary determinant of financial performance. Consequently, in recent years, sustainable investing has notably increased in popularity, as evident in the drastic increase of monetary inflows into the ESG[1] space, reaching $25 trillion in assets under management in 2020 (Global Sustainable Investment Alliance, 2021, p. 10). This shift in focus has necessitated firms and investors alike to further incorporate sustainability considerations (e.g., ESG issues) in their decision-making.

1.1 Background and Motivation

Overall, exploiting available information on firms' actual ESG performance in investment decisions has proven challenging. A commonly used approach by rating agencies is to quantify a firm's sustainability performance by generating ESG ratings, which aim to measure the firm's overall performance on the three pillars of ESG. Nonetheless, there is still no common established practice among the rating providers in measuring sustainability performance. This is partly due to the challenge of agreeing on which ESG issues are deemed most influential. Consequently, there are today a multitude of ESG data providers who apply different methodologies when measuring the sustainability performance of firms. Thus, each rating provider can selectively choose which factors to include and emphasize the ESG pillars differently, arguably adding a subjective element to ESG ratings.

Furthermore, research on the subject finds evidence of a lack of transparency in reporting and regulation practices, and in turn, the problem of asymmetric information, which this provokes, to be more prominent among emerging economies (Diamonte et al., 1996; Duque-Grisales & Aguilera-Caracuel, 2021). Thus, emerging markets are more vulnerable to being influenced by subjectivity in ESG methodologies compared to developed markets. Consequently, one could argue that these differences potentially result in a new

[1] ESG is an abbreviation of the three pillars: Environmental (E), Social (S), and Governance (G).

understanding of how ESG influences financial performance in emerging markets. Hence, we touch upon the inherent differences between the two markets and attempt to broaden the current understanding of ESG on financial performance in emerging economies by applying similar empirical methods as those used in developed markets.

Moreover, the decision to narrow our focus to emerging markets was also influenced by the paradox that emerging economies generally tend to lag behind in terms of ESG issues and have thus been less concerned about sustainability, despite the potential consequences of ESG incidents having a relatively greater negative impact on emerging economies (Mendelsohn et al., 2006). Consequently, emerging economies are presented with untapped potential, as the importance of sustainability concerns has increased globally, thus leading to, e.g., a reduced cost of capital for green investments (Ambec & Lanoie, 2008).

We have observed a void in relevant research regarding the link between ESG and financial performance for emerging markets, as the existing literature predominantly focuses on developed markets. Therefore, while we do not explicitly analyze the effect of the rating methodology inconsistencies, irregularities in measuring reliable sustainability performance and the increased risk and consequences of asymmetric information in emerging markets remain among the motivating factors behind selecting this thesis topic. Additionally, we were intrigued to investigate how this effect proved evident during the recent period of high market volatility as a consequence of the COVID-19 pandemic. Hence, in an attempt to bridge the empirical gap between developed and emerging markets, this thesis aims to consider the effect of ESG performance on stock returns (i.e., financial performance) in emerging markets.

1.2 Methodology

Against this backdrop, we utilize ESG scores as a metric to reflect corporate sustainability performance (i.e., ESG performance) and examine how ESG influences stock returns. The majority of existing research predominantly uses ESG ratings from a single provider. Hence, to expand on the existing research and attempt to capture the overall rating consensus better, we calculate the average ESG score across three different providers (Bloomberg, Refinitiv, and Sustainalytics). For our sample, we use the constituent firms, which comprise the MSCI Emerging Markets Large & Mid Cap Index, where we have

retrieved ESG ratings for 1327 firms. We consider the four years from 2018 through 2021, collecting monthly data for the entire period. To assess the financial performance, we have collected the monthly total returns for each firm.

To conduct our analysis, we group the firms into percentiles based on their overall average ESG ratings. As such, we apply two separate cutoffs, Decile and (Quartile), to select only the 10% (25%) highest- and lowest ESG-rated firms. To capture potential ESG premium or discount and the differences in abnormal returns, we construct a zero-investment portfolio by going long (buy) the high ESG-rated firms and short (sell) the low ESG-rated firms. We rebalance the portfolio annually, as ESG ratings are updated at different intervals for each rating provider. Finally, to account for potential differences in exposure to portfolio risk, we apply the framework of Fama and French (1992) to construct factor portfolios for each respective cutoff. Ultimately, this enables us to investigate the link between ESG ratings and financial performance in emerging markets.

1.3 Results

Analyzing the high- and low portfolios against the market suggests that low ESG-rated firms had relatively higher monthly abnormal returns than high ESG-rated firms. However, we could not establish statistically significant relationships between ESG performance and financial performance for our high-low difference portfolio during our specific sample period, irrespective of model specification and cutoff. Our results add to the inconclusive evidence on the relationship between ESG performance and financial performance in emerging markets but provide valuable insight into the potential driving forces of the high- and low ESG-rated firms' performance during a period of high market volatility.

1.4 Structure

Our thesis is structured as follows. Initially, in Section 2, we present existing research covering an overview of essential terminology, differences in quantifying ESG performance, relevant theories, ESG rating disagreements, and the link between ESG performance and stock performance. Thereafter, in Section 3, we present our research questions and the hypothesis. Next, we present descriptive statistics of our data in Section 4 before outlining the process of sample selection, data preparation, and concerns regarding our data. Next,

Section 5 describes the empirical methodology and the different model specifications used to construct our factor portfolios. Subsequently, we present the results of our analysis in Section 6 before discussing our findings in Section 7. Finally, we round up our thesis and provide a conclusion in Section 8 along with limitations and suggestions for future research.

2 Background

The following section presents background on ESG, financial performance, and relevant economic theories. To better understand why firms should invest time and money into incorporating ESG activities, we first discuss the theoretical framework of shareholder and stakeholder theory. Next, we continue to clarify important terminology and sustainability frameworks before we discuss quantifying ESG performance. Then, we outline ESG disagreement in the context of financial performance and empirical evidence of the link between ESG and financial performance in developed and emerging markets.

2.1 Theoretical Frameworks

ESG considerations have developed to be viewed as important aspects within the theoretical frameworks since they can impact the financial performance and reputation of firms (H. B. Christensen et al., 2019). By incorporating ESG considerations into corporate decision-making, a firm can aim to achieve both its financial and non-financial objectives and generate value for all stakeholders.

2.1.1 Shareholder Theory

However, Friedman (1970) concluded that businesses only have one social responsibility, i.e., to maximize shareholder value (profits) within legal boundaries and fundamental social standards. Consequently, the main social contribution of a company to the public is driven by the ultimate incentive, i.e., profits, made possible by capitalism through undisturbed competition (Kitzmueller & Shimshack, 2012). The Friedman doctrine laid the foundation for shareholder theory and enjoyed widespread support in several academic fields, operating as the dictating corporate objective of many executives in the late 1990s (Smith, 2003).

However, shareholder theory has since been subjected to increased criticism, especially after several corporate auditing discrepancies were uncovered in the early 2000s, e.g., the Enron scandal. Thus, Friedman's views have been criticized for lack of nuance and for motivating firms to provide financial returns at all costs, even though Friedman explicitly states that each firm needs to abide by the current laws and regulations. As a result of this

misconception, there is an important distinction to be made between what shareholder theory is and its implications. Critics are quick to point out evidence in the wake of the legal repercussions of financial scandals, referring to managers' careless attitude toward making profits at all costs. It is a prevalent practice that executives are largely incentivized through financial compensation, such as stock options and other incentive schemes. Nevertheless, critics claim that such incentives motivate executives to respond accordingly and implement any measures necessary to ultimately maximize profits (Smith, 2003). However, one could argue this to be a false narrative, as Friedman (1970) clearly states that the pursuit of profits should "stay within the rules of the game" and "without deception fraud". According to Smith (2003), another common misconception is that shareholder theory encourages short-term managerial thinking at the expense of the long run. Meanwhile, Danielson et al. (2008) argues this perception to be misguided because shareholder maximization is inherently a long-term goal, where firms must maximize all future cash flows.

2.1.2 Stakeholder Theory

In the widely cited article, *Stockholders and Stakeholder: A new perspective on Corporate Governance*, Freeman and Reed (1983) laid the foundation of what is today known as stakeholder theory. Stakeholder theory asserts that an organization's success depends on how well it manages relationships with key groups, such as customers, employees, suppliers, communities, financiers, and others, that can affect the realization of its purpose (Freeman & Phillips, 2002, p. 333). As a result, scholars began questioning managers' attitudes toward sustainability when the ultimate objective was maximizing shareholder wealth. Thus, since its introduction, stakeholder theory has become an integral part of further study in several academic fields.

While shareholder theory proposes that shareholders are the only group with a moral claim on the corporation, stakeholder theory contends that corporations' responsibility extends beyond shareholders, i.e., stakeholders, to incorporate causes that benefit society as a whole. Freeman (2001) presents two main arguments as to why shareholder theory offers an overly simplistic view: Firstly, society's continuously amended laws and regulations have significantly restricted the pursuit of profits. Secondly, economic externalities, moral

hazards in the form of "pass along costs"[2] or monopolies are all realities which the organization needs to consider. The former argument, labeled the "legal argument", suggests that shareholder theory is practically useless. Whereas the latter argument, labeled the "economic argument", states that shareholder theory neglects to account for established economic realities, thus invalidating shareholder theory.

Additionally, stakeholder theory argues that corporations can either harm (e.g., global warming) or benefit (e.g., public goods) the society it operates in. As a result, firms must address these demands when making corporate decisions or risk facing hostile confrontations from affected groups. If such requests are not respected, it could potentially lead to diminishing returns due to lawsuits, demonstrations, etc. (Freeman & Reed, 1983; Ruf et al., 2001). Conversely, suppose the preferences of stakeholders align with those of the firm. In that case, a common justification is that they could positively enhance stakeholders' firm reputation and loyalty, with the potential to affect a firm's financial performance (H. B. Christensen et al., 2019, p. 48).

2.2 Sustainability Frameworks

Given the potential environmental threats that follow inconsiderate investing, several frameworks have been introduced to further accelerate companies' sustainability practices. Moreover, when it comes to ESG and sustainability, terms such as ESG, CSR, and SRI cover similar sustainability-related issues. Thus, while we present the underlying meaning of each term, we consider research within each respective area to be reflected by the overall ESG performance. We assume that research on these issues can be used interchangeably when analyzing ESG performance. As such, the following parts of this thesis predominantly use ESG as an overarching term when referring to sustainability-related issues.

2.2.1 Corporate Social Responsibility

Corporate Social Responsibility (CSR) is the most predominant term used in the literature (H. B. Christensen et al., 2019; Huang & Watson, 2015). In essence, CSR can be defined as voluntary corporate behavior that improves social welfare and exceeds legal and regulatory requirements (Kitzmueller & Shimshack, 2012; McWilliams & Siegel, 2001). This means

[2]Externality related costs which affects ("passed along to") third parties.

businesses have taken on more responsibility than what is strictly required to meet the growing social demands of stakeholders.

However, it is not entirely apparent that firms should voluntarily engage in CSR activities, as this may exhaust firm resources, reduce profits and shareholder wealth (Huang & Watson, 2015; Walley & Whitehead, 1994). Ideally, by engaging in CSR, firms experience improved reputation through increased demand, customer loyalty, employee retention, etc. Similarly, CSR has been found to help repair reputational damage following, e.g., negative earnings restatements (Chakravarthy et al., 2014). On the other hand, the implementation of CSR could also be propelled by strategic incentives or induced by markets. To the former, Kitzmueller and Shimshack (2012, p. 26) argue that firms might strategically hedge against potential regulations and related costs adjustments through CSR over-compliance. By doing so, firms are better equipped to preserve their competitiveness when facing potential new regulations but also discourage future regulations by implying that markets already provide adequate incentives. It seems obvious, however, that firms are not exclusively motivated by profit-maximizing activities as suggested by neoclassical economics but also incentivized to some degree by CSR presented through shareholder theory.

2.2.2 Socially Responsible Investing

Another abbreviation in the sustainability terminology that accounts for sustainability in the investment process has been labeled socially responsible investing (SRI). However, within the academic literature, there does not seem to be a universal agreement on what the term implies for investors. A possible explanation could be that the main focus of these studies has been on the impact of SRI on financial performance rather than SRI itself (Berry & Junkus, 2013). In a review of the development of socially responsible investing, Camilleri (2020) argues that investors are intrinsically attracted by financial instruments that yield a financial return on their investments. Nevertheless, there has also been a rise in investors who integrate their personal values into investment decisions. Since personal values differ for each investor, SRI will inherently not be similar for everyone. Recently, many investors have been inspired by the potential to implement social and environmental goals into their investment decision to justify that their investment decisions address societal and community deficits. Therefore, responsible investors expect encouraging and

visible results from their capital allocations to the environment and community. Thus, one can contemplate SRI as a complementary investment strategy with the primary goal of benefiting investors with financial returns and an underlying focus to simultaneously improve social -and environmental welfare.

2.2.3 Environmental, Social, and Governance

ESG is another framework that has emerged as an essential pillar of CSR (Eccles et al., 2013) and aims to encapsulate a firm's sustainability efforts. Over the last 25 years, the amount of companies that measure and disclose ESG data has increased exponentially due to the importance of environmental issues (Amel-Zadeh & Serafeim, 2018). Correspondingly, a growing amount of investors incorporate ESG issues into their decision-making and investment research. According to Forum for Sustainable and Responsible Investment (US SIF, 2020, p. 9), one-third of all investment assets under professional management in the United States are either considering ESG issues or filing shareholder resolutions on ESG issues at publicly traded companies. This emphasizes that environmental and social awareness has become a societal focal point for corporations and investors (Amel-Zadeh & Serafeim, 2018). Importantly, responsible investors want to receive a financial return that reflects the market return and makes a positive social outcome (return). While the two are not necessarily complementary, it reveals the potential existence of other non-pecuniary incentives that could drive current investment allocations.

2.3 Differences in Quantifying ESG Performance

Asset owners and portfolio managers have increasingly focused on incorporating ESG considerations into their investment strategies. However, investors lack guidance on incorporating these considerations in portfolio decisions. This is further complicated by the diverging views among academics and practitioners about the effect of ESG on portfolio performance (Pedersen et al., 2021). Accordingly, the actual corporate sustainability and overall ESG performance of firms are, in practice, difficult to establish objectively without firm-specific insight.

2.3.1 ESG Ratings

As such, professional rating agencies have developed several metrics in an attempt to quantify ESG performance across the environmental, social, and governance dimensions. Rating agencies collect information from numerous sources, culminating in an overall ESG score that investors consider when making responsible investment decisions. As stated by Duque-Grisales and Aguilera-Caracuel (2021), the responsible investor evaluates this decision based on three ESG dimensions or pillars. Firstly, a company's environmental score is a weighted score of strengths and weaknesses on indicators related to i) emissions reduction, ii) product innovation, and iii) resource consumption reduction. Thus, a company that is able to control and prevent pollution will typically receive a high environmental score. Secondly, social performance refers to managing primary stakeholders, i.e., employees, customers, and the community it operates in. The social pillar addresses firm policies regarding working conditions, workplace diversity, training and labor rights, employee and customer satisfaction, business relationships, and other issues relevant to interested parties. Thirdly, the governance score measures to what extent a company's systems and processes ensure that its members and executives operate in the best interest of its shareholders. Professional investors frequently scrutinize the governance pillar because of its close relationship with determining the quality of management (van Duuren et al., 2016). Audit and board independence, completion of sustainability reports, information transparency, corporate ethics, and minority shareholders' rights are all examples of corporate governance-related issues (Duque-Grisales & Aguilera-Caracuel, 2021; Miralles-Quirós et al., 2018).

Quantifying actual firm ESG performance across these three pillars helps to ensure that businesses are held accountable to shareholders. Moreover, the increased transparency and insight into firm ESG performance provided by ESG scores, beyond what firms are legally obliged to report, are also valued by external stakeholders. Naturally, investors will appreciate accurate reporting practices due to the limited insight into an organization's activities they provide. The fact that an increasing number of companies unveil their ESG activities indicates that adjustments in business strategy have been incorporated to stress the importance of investor relations through public communication.

2.4 ESG Disagreement

Relevant research remains divided on the effect of ESG on portfolio returns. Yet, ESG considerations, most commonly materialized through ESG scores, continue to influence investor decisions. Accordingly, this constitutes a major concern regarding the methodology used to generate these metrics and the following disagreement among providers. According to 2020 estimates from the Global Sustainable Investment Alliance (2021, p. 10), the most prominent sustainable investment strategy globally is ESG integration, with a combined $25.2 trillion in assets under management employing ESG factors into financial analysis. Moreover, as reported in late 2021, more than 100 different ESG data providers allow investors to screen selected companies for ESG performance (Zehetmayr & Brandau, 2021). Such services could thus prove invaluable for investors in detecting problematic ESG issues.

2.4.1 Differences in Rating Methodologies

However, despite how providers primarily base their ESG ratings on the same three environmental, social, and governance pillars, there has been an emerging problem directed at the disagreement or divergence of ratings across different rating providers for the same firm (Gibson Brandon et al., 2021). Moreover, all ESG providers apply their own subjective methodology and theoretical biases, which might complicate investors' ability to draw meaningful conclusions about their investment decision (D. M. Christensen et al., 2022).

Given the existing diversity of ESG rating methodologies, this issue has also sparked significant academic interest among researchers. In their paper, Berg et al. (2019) investigate what drives the disagreement of sustainability ratings. The paper highlights three sources of ESG rating divergence; i) raters use different categories (scope divergence); ii) raters measure identical categories differently (measurement divergence), and iii) raters assign different weights to different categories (weight divergence). They emphasize that it is challenging to interpret the differences between two ESG ratings because of the interplay between scope, measurement, and weight divergence. The paper argues that most discrepancies could be linked to scope and measurement divergence. For that particular reason, it becomes challenging to resolve. Furthermore, they discover that

measurement divergence is partly driven by a substantial *rater effect*, also known as a halo effect. This implies a bias, where a firm receiving a high score in one category is more likely to receive high scores in every other category from that same rating provider. Consequently, the abovementioned findings should urge investors not to become overly reliant on a single rating agency.

2.4.2 Asymmetry in ESG Disclosure

A working paper by D. M. Christensen et al. (2022) examines whether a firm's ESG disclosure helps explain discrepancies across rating agencies. Even after controlling for firm fixed effects, the authors find that greater ESG disclosure leads to higher ESG rating disagreement. Their paper argues that due to the subjective nature of ESG information, relating to the different methodologies used by rating providers, higher disclosure and increased transparency are associated with higher disagreement between providers. This represents a counterintuitive phenomenon, where increased disclosure and reduced information asymmetry actually widen the possibilities for different interpretations. In addition, higher disclosure raises the possibility that ESG providers can use various metrics to assess a company's performance on the same issue, which could result in even more rating disagreements. Conversely, in the absence of ESG disclosure, rating agencies are more likely to agree as they are more inclined to view the lack of disclosure as an inferior feature and thereby assign the firm a low score (D. M. Christensen et al., 2022).

Additionally, Chatterji and Toffel (2010) argues that ESG scores can mitigate the adverse selection problem when faced with information asymmetry and incomplete information about a firm's ESG performance. As such, ESG scores serve an intermediary function by helping investors and stakeholders to take ESG considerations into portfolio decisions.

2.5 ESG Disagreement and Financial Performance

Although previous literature has primarily focused on the reasons *why* ESG ratings disagree, Gibson Brandon et al. (2021) study the *impact* of ESG rating disagreement on stock returns. Their findings suggest a positive correlation between stock returns and ESG rating disagreement, which means that firms with more significant disagreement tend to have higher stock returns. Additionally, Gibson Brandon et al. (2021) find

that disagreements about the environmental dimension are the primary driver of this relationship. As a result, the authors conclude that this is coherent with the idea that risk-averse investors recognize a dispersed ESG performance of a given firm as an additional source of risk (or uncertainty) that demands a separate risk premium to hold the stock. In addition, their results produce important practical implications for investors who aim to optimize both financial performance and responsible investment strategies.

Moreover, implementing sustainable firm strategies is likely to be costly, resulting in a sacrifice of short-term earnings for long-term outcomes (Starks et al., 2017). Conversely, Porter and van der Linde (1995) argue that properly designed environmental strategies have the potential to spur efficiency and technology innovation, resulting in improved cost savings and increased profitability. However, from an investor's viewpoint, the benefits of investing in firms that implement ESG strategies will depend on the investor's time horizon. As such, the presence of asymmetric information may lead to diverging evaluation of ESG projects among long- and short-term investors (Starks et al., 2017).

Overall, ESG rating disagreement is important because of the growing amount of investors that consider ESG scores in their investment decisions and has consequently been subject to increased attention by policymakers, academic research, and the financial press. Generally, ratings are provided to guide investors in their investment process. However, ESG discrepancies appear to have had the opposite effect, at least to some extent. This becomes even more pronounced for investors when various providers' evaluations of the same firm have extensive disparities. While acknowledging that different raters come to different conclusions on which categories should be given the most weight, it still seems unlikely to settle the ESG debate until it is possible to establish common standards or frameworks for what constitutes good and bad ESG performance (Berg et al., 2019; D. M. Christensen et al., 2022).

However, it is important to note that while the importance of ESG rating disagreements should not be understated, we do not *explicitly* analyze the actual effects of this disagreement but rather attempt to capture the overall rating consensus by using multiple ratings in our analysis.

2.6 ESG Scores and Financial Performance

The relationship between ESG performance and corporate financial performance (CFP) has been subject to extensive research since the beginning of the 1970s (Friede et al., 2015). However, despite the vast literature on the subject, researchers have yet to reach a general agreement on the implications of this link. Moreover, new ambiguity on the subject frequently occurs as more research emerges, leading to fragmented knowledge and inconclusive results (Atz et al., 2022; Friede et al., 2015; Griffin & Mahon, 1997).

2.6.1 Positive Empirical Evidence

Kempf and Osthoff (2007) investigated whether implementing a long-short trading strategy based on SRI ratings from KLD Research & Analytics (now MSCI) would yield higher financial returns for investors. Specifically, the authors hypothesized that going long firms with high SRI ratings and short firms with low SRI ratings, included in the S&P 500 and Domini 400 Social Index, would lead to financial outperformance. Indeed, for the period 1992-2004, following the long-short strategy yielded annual abnormal returns of up to 8.7%. These results held even after adjusting for reasonable transaction costs. Similarly, Eccles et al. (2014) studied 180 U.S. listed companies, where half were classified as high sustainability companies, while the other half were classified as low sustainability companies. Consistent with their expectations, Eccles et al. (2014) found that high sustainability companies outperform low sustainability companies concerning stock market and accounting performance, such as ROE or ROA, from 1993 to 2010. These findings are further substantiated by a previous study of Eccles et al. (2013), which suggests that the development of sustainable strategies plays a significant role in how ESG scores impact financial performance.

In their meta-analysis, Friede et al. (2015) investigated aggregated evidence from over 2000 empirical studies since the beginning of the 1970s. Interestingly, they documented that approximately 90% of the included studies find a non-negative ESG-CFP relation that also appears stable over time. Moreover, based on correlation and distribution factors, their results indicate that ESG criteria and CFP, on average, are positively correlated. Their analysis also covers research on potential differences in the ESG-CFP relation across geographical regions. According to their analysis, two main patterns in the regional data

stand out. First, developed markets reveal a smaller share of positive results excluding North-America. For example, in developed Europe, the share of positive results amounts to 26.1%. Meanwhile, developed Asia/Australia has a positive share of 33.3% – although forming the largest share of negatives at 14.3%. However, these results are possibly biased by the underlying studies comprising a larger share of portfolio-based studies. Secondly, the emerging market's sample indicates a significantly higher share of positive outcomes relative to developed markets, amounting to 65.4%. This percentage becomes even higher at 70.8% when omitting portfolio-based studies from the sample.

These findings are broadly consistent with another meta-analysis of more than 200 academic studies conducted by Clark et al. (2015). According to their report, 80% of the studies that were reviewed demonstrate that prudent sustainability practices positively impact stock price performance. In addition, 88% of the research shows a positive correlation between the operational performance of firms and ESG practices.

2.6.2 Negative Empirical Evidence

In contrast, some results indicate the opposite. For example, in an analysis, Vance (1975) compared the performance of firms classified as either having high or low levels of CSR. He found that firms with low levels of CSR outperformed firms with high levels of social responsibility. Furthermore, he found a negative relationship between corporate social responsibility and financial performance. Thus, he concluded that socially responsible firms were not desirable as investment objects. Analogously, a more recent study by Sargis and Wang (2020) found that investors would have slightly underperformed in holding better ESG securities if they limited their holdings to U.S. and Canadian securities. The authors imply that investors pay higher prices for good ESG companies as there might exist a premium in return, making investors more inclined to hold companies with bad ESG practices. Moreover, investors should demand to receive this premium as compensation for any ESG-related risk. Nevertheless, the study found that investing in ESG on a global scale did not involve any risk/reward trade-offs.

2.6.3 Inconclusive Empirical Evidence

Lastly, there is also a body of literature where results are inconclusive. For instance, Atz et al. (2022) surveyed 27 meta-reviews and 1,141 peer-reviewed papers published between 2015 and 2020. Overall, their findings indicate that, on average, financial performance from ESG investing has been indistinguishable from traditional investing. Similar results have been made, where Revelli and Viviani (2015) concluded that pursuing SRI does not add financial costs or benefits with respect to traditional investments. This is also consistent with Renneboog et al. (2008), who documented no significant relationship between ESG and stock returns.

2.7 Empirical Evidence from Emerging Markets

Most studies on the link between CFP and ESG performance have been centered around markets in developed economies, mainly in the U.S., European-, and some Asian-pacific countries (Auer, 2016; Daugaard, 2020; Wang et al., 2022). The main culprit is that reliable data were largely unavailable until more recently, thus leaving emerging markets an underresearched area relative to developed markets. However, recent research indicates a particularly positive ESG-CFP relationship in emerging Markets. For instance, Friede et al. (2015) reported a 65.4% higher share of positive results in emerging markets than in developed ones.

2.7.1 Evidence from BRICS-Countries

Garcia et al. (2017) examined businesses from Brazil, Russia, India, China, and South Africa (hereafter BRICS) to determine whether financial performance is associated with superior ESG performance. In particular, the study analyzed ESG performance in sensitive industries usually characterized by social taboos, moral debates, and political pressures. Data from 365 BRICS-listed companies were collected and divided into four different ESG performance metrics: i) overall ESG performance, ii) environmental performance, iii) social performance, and iv) corporate governance performance. Even after controlling for firm size- and country effects, the results indicate that firms in sensitive industries, or those being more likely to cause harm to society, present superior environmental performance. More specifically, they report a negative association of financial performance

with environmental performance. Given that sensitive industries consist of firms with the greatest environmental impact, these findings could support the idea that they are more consistent in disclosing ESG practices to legitimize their operations. By extension, greater disclosure is necessary for riskier and more aggressive firms (i.e., sensitive industries) to minimize informational asymmetry and lower the cost of capital of the firm (Garcia et al., 2017; Mendes-Da-Silva et al., 2014).

A more recent study by Duque-Grisales and Aguilera-Caracuel (2021) analyzed the link between ESG scores and the financial performance of multinationals in Latin America (i.e., multilatinas) during 2011-2015, listed as emerging markets in Brazil, Chile, Colombia, Mexico, and Peru. The results imply that a high ESG score leads to worse financial performance for multilatinas. Furthermore, multilatinas are frequently characterized by lacking financial flexibility due to the scarcity of resources. As such, managers tend to pursue more profitable operational activities rather than prioritizing ESG initiatives since they are thought too costly (Sharma, 2000). Yet, managers are less concerned with short-term expenses if the firm has adequate financial flexibility. Accordingly, multilatinas are increasingly likely to support ESG initiatives in response to stakeholder pressure (Duque-Grisales & Aguilera-Caracuel, 2021). Moreover, multilatinas systematically differ from developed markets concerning social and cultural aspects. Consistent with these findings, Diamonte et al. (1996) argues that political and institutional differences generally represent a more significant determinant of stock returns in emerging markets relative to developed markets.

2.7.2 Contradicting Evidence

Overall, it is important to note that several possible explanations for the contradicting results between ESG and financial performance are observed in the literature. Firstly, researchers differ in their use of metrics to capture financial performance. Naturally, regarding financial performance, several applicable accounting- and financial metrics (i.e., ROA, ROE, EVA, FCF, etc.) exist and would consequently provide different results. Secondly, the research varies partly because of the numerous rating methodologies used by different rating providers and differences between geographical regions. As initially discussed, ESG is a broad term with various interpretations that allow rating agencies to apply different theoretical frameworks to determine sustainability performance. Thus,

introducing a bias where ratings are, at least to some degree, exposed to subjectivity from the rating providers. Another bias is introduced due to studies predominately focusing on American or European markets, leaving other regions an underresearched part of the literature. Ideally, some of the results are still generalizable based on similarities in study design. However, there are institutional differences between regions that cannot be overlooked. Hence, this calls for thorough consideration before reaching definite conclusions.

3 Hypothesis Development

This section presents the overarching research question, supplementary research question, and hypothesis. In addition, we discuss the motivation behind our decision to analyze emerging markets and clarify the underlying assumptions of our research questions.

Compared to developed markets, emerging markets remain a relatively underresearched area. Moreover, emerging markets are systematically different in terms of cultural, political, and institutional-related aspects (Diamonte et al., 1996; Duque-Grisales & Aguilera-Caracuel, 2021), and the implications of these discrepancies remain the primary motivation for our choice of market. Consequently, by investigating the link between ESG performance and financial performance in emerging markets, our analysis contributes to the fragmented and divided research on the area.

3.1 Overarching Research Question

In this thesis, we attempt to answer whether there exists a link between ESG performance and stock returns in emerging markets by answering the following research question:

How does ESG performance impact the expected stock returns of firms in emerging markets?

To this extent, we seek to investigate whether there exists a link between ESG performance and stock returns and the size of this effect. Intuitively, we expect the associated risk to be higher with low ESG-rated firms and thus resulting in investors requiring increased compensation in terms of higher expected returns. Accordingly, we aim to identify the influence of ESG performance on firms' financial performance and whether it helps to explain emerging market stock returns variation. If present, we assume this risk to be particularly pronounced for emerging markets relative to developed markets due to their increased sensitivity to ESG concerns.

3.1.1 Supplementary Research Question

To expand our understanding of the overarching research question, we are interested in examining how this effect is influenced by the relative difference in ESG performance between the high- and low ESG-rated firms' restrictiveness when defining what constitutes

a high- and low ESG performer.

How does the restrictiveness of defining what constitutes high- and low ESG performance influence stock returns?

We expect that the restrictiveness when deciding what determines a high- and low ESG-rated firm impacts the size of the premium investors pay for holding the firm. Hence, we expect the premium to increase (decrease) when applying a more (less) restrictive definition of a high (low) ESG-rated firm, i.e., holding firms with relatively better (worse) ESG performance results in a higher (lower) premium paid.

However, it is important to note that we will not directly analyze the difference between Decile and Quartile, as we primarily limit the focus of our analysis to examining the differences between high- and low ESG-rated firms. Hence, we only discuss the relative implications of applying the two definitions of what constitutes a high (or low) ESG performer. Thus, we aim to use the two cutoffs to supplement our understanding of the overarching research question.

3.1.2 Hypothesis

While contradicting evidence from the relevant research makes it unclear whether we will identify a positive, neutral or negative link, we have developed the following hypotheses:

Hypothesis: *Investors pay a premium for holding firms with good ESG practices in emerging markets, in the sense that the expected return of high ESG-rated firms is lower than low ESG-rated firms.*

We expect high ESG-rated firms to be associated with lower risk and, thus, lower expected returns. On the contrary, we expect low ESG-rated firms to be associated with higher risk; therefore, investors need to be compensated by higher expected returns. Accordingly, investors holding high ESG-rated firms pay a premium equal to the opportunity cost by foregoing higher returns when holding low ESG-rated firms.

To answer our research questions, we seek to examine and quantify the intuition described by our hypotheses. We address each hypothesis by constructing a high (low) portfolio that goes long (short) in high (low) ESG performers. Furthermore, we address the differences between the high- and low portfolios by generating a zero-investment portfolio. As a

result, we assume that stock performance reflects all available information, i.e., the firms' overall corporate financial performance, under the efficient market hypothesis (EMH). Moreover, we assume that overall firm ESG performance is reflected by the ESG scores extracted from prominent rating providers.

4 Data

This section describes the process of selecting and preparing the data used in our analysis. First, we identify the data sources used before elaborating on our choice of variables and providing descriptive statistics of our sample data. Next, we present important remarks concerning how the different portfolios and variables were constructed and calculated. Then, we outline the main steps of the data preparation, highlighting the necessary modifications made to the data. Lastly, we round up this chapter with a discussion of the concerns and limitations regarding our data.

4.1 ESG Ratings and Rating Providers

We use ESG scores in our analysis to quantify firm ESG performance. However, as elaborated on in Section 2.3, the scoring methodology for quantifying ESG performance varies between different ESG rating providers. Hence, consistent with the discussion in Section 2.4.1, to avoid introducing selection bias by only using a single provider and to better reflect the consensus of each firm's ESG performance, we collected ESG scores from three different rating providers: *Bloomberg*, *Refinitiv*, and *Sustainalytics*.

4.1.1 Bloomberg

Bloomberg and their associated third-party sources collect, verify and provide ESG data on more than 2000 ESG fields and scores for more than 14 000 companies worldwide (Bloomberg, 2022a). Bloomberg's ESG data covers ESG metrics for several sectors and countries dating back to 2006, which allows investors to analyze firm performance over time. The collected data is based on public CSR reports, annual reports, company websites, continuous communications with firms, and a survey that inquires about corporate information directly (Basar, 2021; Bloomberg, 2022b). The Bloomberg proprietary overall ESG score encompasses firms' absolute and relative performance across the three pillars of corporate environmental, social, and governance performance. Bloomberg's proprietary overall ESG score ranges from worst (0) to best (10).

4.1.2 Refinitiv

Refinitiv, previously a part of Thomson Reuters' Financial & Risk business, was formed in October 2018 after Blackstone Group acquired 55% of the majority shares from Thomson Reuters. Subsequently, in August 2019, Thomson Reuters and Blackstone struck an agreement with the London Stock Exchange Group (LSEG) to sell Refinitiv (Nasdaq, 2019, 2021).

Refinitiv is one of the world's leading financial market and infrastructure data providers. Hereunder, Refinitiv supplies information and insights into more than 40 000 institutions in 190 countries (Refinitiv, 2022a). To better fit investors' interest in sustainable investment decisions, Refinitiv has since 2002 incorporated ESG services, which currently cover 85% of the world's market capitalization across more than 630 unique ESG metrics. The ESG scores are based on how ESG factors performed in relation to a given firm's industry (Refinitiv, 2022c). Therefore, to ensure objective and transparent measures of ESG performance, Refinitiv offers ESG data contingent on publicly available sources (Refinitiv, 2022b). As a result, we used Refinitiv's score, labeled *ESG Score*, in our research[3], which is an overall assessment of the organization derived from self-reported data across the three environmental, social, and governance pillars. Refinitiv implements a percentile rank methodology across ten different categories, which results in three pillar scores that are a relative sum of the category weights. Ultimately, the pillar weights are normalized to a percentage score between 0 (worst) and 100 (best) (Refinitiv, 2022c).

4.1.3 Morningstar Direct - Sustainalytics

To better advance their sustainability goals, on April 21, 2020, Morningstar Inc. announced they had reached an agreement to acquire Sustainalytics, a specialist in ESG rating and research (Nasdaq, 2020). Sustainalytics measures the extent to which a company's economic value is at risk due to its ESG factors (Sustainalytics, 2021), as opposed to Bloomberg's and Refinitiv's ESG ratings, which primarily focus on ESG performance. By quantifying the amount of an organization's unmanaged risk, this score has been

[3]Refinitiv also supplies the *ESG Combined Score*, which is based on a company's ESG score while also accounting for controversies associated with that company in a certain reporting period. However, we deemed the *ESG score* more compatible for our purpose since the use of *ESG Combined Score* could introduce biases that were not equivalent to the score of the other two providers.

named Sustainalytics' *ESG Risk Rating*, which aims to shed light on firm-level ESG risk. Furthermore, over 14 000 firms and major global indices are covered by Sustainalytics' ESG Risk Rating (Sustainalytics, 2022). Three building blocks comprise the ESG Risk Rating: i) Corporate Governance, ii) Material ESG Issues, and iii) Idiosyncratic ESG Issues. These provide the foundation for a company's overall rating, which is then assigned into five ESG risk categories that could affect a company's economic value, i.e., negligible, low, medium, high and severe. Thus, the final score is analogous to a numerical scale ranging from 0 (negligible risk) to 100 (most severe risk), which is suitable for comparison between industry peers (Sustainalytics, 2022).

4.2 Financial Metrics

To have confidence in our results, the financial metrics need to reflect the financial performance of firms accurately. Therefore, our financial metrics consist of firms' *total return* and *market capitalization*. Additionally, we normalize the denoted currency of returns in the cross-section of our sample by only collecting data in U.S. dollars. As such, we make it easier to reproduce our findings and the comparability to relevant research.

4.2.1 Total Return

We collected monthly total return data[4] from the Morningstar Direct platform. By taking the change in monthly Net Asset Value (NAV), reinvesting all income and capital-gains distributions for that month, and dividing by the starting NAV, yields the total return with a monthly frequency. The monthly reinvestment of daily payoffs is done using the actual reinvestment NAV (Morningstar, 2022). Moreover, because all return data is extracted in U.S. dollars, we avoid the need to convert exchange rates manually.

4.2.2 Market Capitalization

Firm market capitalization was collected from Refinitiv, labeled *Market Cap*. The Market Cap represents the sum of the firm's market value for all relevant issue level share types and is calculated by multiplying the outstanding shares by the latest close price, as presented in Equation (4.1):

[4]The total return is calculated using the Adjusted Closing Price and is denoted in U.S. dollars.

$$Market\,Cap_{i,t} = Price_{i,t} \times Outstanding\,Shares_{i,t}, \tag{4.1}$$

where $Market\,Cap_{i,t}$ denotes the firm market capitalization for firm i, $Price_{i,t}$ denotes the latest adjusted closing price for firm i, and $Oustanding\,Shares_{i,t}$ denotes the total number of outstanding shares for firm i, at time t.

4.3 Sample Selection

When conducting ESG-related research, the sample selection is generally exposed to the restrictions imposed by the ESG rating availability, which exists both in the cross-section and the time series of our sample (Gibson Brandon et al., 2021). Thus, to test our hypothesis on a representative and homogeneous sample, we extracted ESG ratings from 1327 (out of the total of 1387) constituents of the MSCI Emerging Markets Index. This index consists of 1386 constituents[5], and captures large- and mid-capitalization representation across 24 emerging markets countries[6]. It is estimated to cover approximately 85% of the free float-adjusted market capitalization in each country (MSCI, 2022). Moreover, by limiting our sample to only include large- and mid-capitalization firms, we narrow the focus of our analysis to only consider the effects of ESG performance on the most influential firms, in terms of market value, in emerging markets. However, this is contingent on the assumption that firm market capitalization accurately reflects firm influence in its respective market.

4.4 Screening

To maximize the number of available ESG ratings, we restrict the time period of our sample to the years 2018 through 2021. This period includes relatively normal market development in the years leading up to 2020 and the high market turmoil that followed after the COVID-19 pandemic in the beginning of 2020. Consequently, our sample allows

[5]Our sample consists of slightly fewer constituents than the MSCI index because the complete list of MSCI Emerging Markets Index constituents was unavailable to us through any database at our disposal. Accordingly, we extracted the constituents of the Bloomberg Emerging Markets Large & Mid Cap Total Return Index, which use the MSCI index as its benchmark index.

[6]The included emerging markets countries are: Brazil, Chile, China, Colombia, Czech Republic, Egypt, Greece, Hungary, India, Indonesia, South-Korea, Kuwait, Malaysia, Mexico, Peru, Philippines, Poland, Qatar, Saudi Arabia, South Africa, Taiwan, Thailand, Turkey, and United Arab Emirates.

us to consider the effect of ESG performance in times of high market uncertainty and under normal market development[7].

Generally, political and institutional differences have historically implied a lower degree of transparency and ESG disclosure for emerging markets (Diamonte et al., 1996), relative to developed markets. Moreover, ESG score availability is more scarce for firms in emerging markets relative to developed markets and differs widely between all providers. Therefore, to overcome the limited availability of scores and to better reflect the consensus of each firm's ESG performance, we calculated an ESG score for each firm using the mean of ratings from all three providers. As discussed in Section 2.4.1, by using the calculated mean ESG score, we attempt to even out any differences between each respective rating provider, thus reducing the potential introduction of subjective bias from a single rating provider. In addition, we also increase the number of available ratings for our sample when some firms lack ratings from any provider, as the providers complement one another. However, by not requiring ESG ratings from at least two providers, we do not exclude the possibility of including a firm that is only rated by one of the providers in the portfolio. Moreover, to avoid restricting our sample size, we neither require the firms in our sample to have ratings for all months in a period nor all periods.

4.5 Data Preparation

As none of the ESG ratings from Bloomberg, Refinitiv, and Sustainalytics had the same format in which their ratings were denoted, we had to modify the data format. Refinitiv denotes its ESG ratings as a score between 0-100, where 100 is the best. To denote the ESG ratings from Bloomberg in the same 0-100 format, the rating was multiplied by 10 to increase the scale of the rating to 100.

The ESG risk ratings from Sustainalytics denote the firm's exposure to ESG risks, with a rating scale of 0-100, where 100 is the worst. Thus, the Sustainalytics ESG risk ratings of Sustainalytics were amended so that 100 reflected the best score by subtracting each ESG risk rating of firm i, at time t from 100 as shown in Equation (4.2):

[7]Normal market development is here referring to the period between 2018-2020 and the destructive effects which followed COVID-19.

$$Score_{i,t} = 100 - Risk\ Score_{i,t}. \tag{4.2}$$

4.6 Descriptive Statistics of Sample

In the following, we provide descriptive statistics of the collected ESG scores of our sample from each respective rating provider to present an overview of the main attributes of our sample data.

Table 4.1: Descriptive Statistics: Bloomberg ESG Scores

This table presents descriptive statistics of Bloomberg's available ESG scores for firms in our sample. The observation period spans the years 2018 through 2021, with a monthly frequency, and the statistics are denoted as yearly averages. The table describes the total number of rated firms N(Firms), mean ESG scores (Mean), standard deviation (SD), minimum- (Min) and maximum (Max) values, and the 25^{th}, 50^{th}, and 75^{th} percentiles.

	2018	**2019**	**2020**	**2021**
N(Firms)	165.00	170.00	179.00	187.00
Mean	27.93	28.82	30.20	30.92
SD	11.16	11.00	10.82	10.80
Min	8.20	8.80	9.30	8.10
25%	19.90	20.07	22.25	22.45
50%	25.90	28.60	30.70	31.30
75%	35.05	34.86	36.05	36.95
Max	64.90	56.80	59.20	59.50

We collected and transformed the proprietary Bloomberg overall ESG score, ranging from worst (0) to best (100)[8] to evaluate a company's aggregated ESG performance. Descriptive statistics of Bloomberg's ESG scores are presented in Table 4.1. As the number of observations in Table 4.1 demonstrates, Bloomberg covers only a limited number of firms compared to the total sample size (1327) of our benchmark index. However, from 2018-2021, there was a slight increase in Bloomberg's emerging market firm coverage. This could reflect the growing focus and demand from stakeholders on ESG transparency. Moreover, Table 4.1 illustrates the ESG scores from Bloomberg for each year, ranging from 0 to 100. For instance, the mean ESG score for 2021 is 30.92, indicating that, on average, Bloomberg's respective methodology does not rate firms highly. This is

[8]The default range of Bloomberg proprietary ESG scores is from worst (0) to best (10). However, as elaborated on in Section 4.5, we normalize the format of Bloomberg's ESG ratings to match the format of the other rating providers.

further substantiated by the max ESG score for 2021 of 59.50, which implies that the best ESG performer in our Bloomberg sample still has the potential to score 40 points better.

Table 4.2: Descriptive Statistics: Refinitiv ESG Scores

This table presents descriptive statistics of Refinitiv's available ESG scores for firms in our sample. The observation period spans the years 2018 through 2021, with a monthly frequency, and the statistics are denoted as yearly averages. The table describes the total number of rated firms N(Firms), mean ESG scores (Mean), standard deviation (SD), minimum- (Min) and maximum (Max) values, and the 25^{th}, 50^{th}, and 75^{th} percentiles.

	2018	**2019**	**2020**	**2021**
N(Firms)	802.00	899.00	1081.00	1178.00
Mean	47.88	48.64	47.47	49.42
SD	20.25	20.73	21.12	20.80
Min	1.10	0.80	0.72	0.96
25%	32.61	33.85	30.33	33.93
50%	48.53	49.37	48.14	50.80
75%	62.73	64.46	63.90	65.63
Max	91.60	92.80	92.26	92.97

As Table 4.2 illustrates, Refinitiv has considerably more ESG observations for the relevant period than Bloomberg. By 2021, Refinitiv had provided 1178 ESG scores out of 1327 firms. In percentage terms, Refinitiv has a substantial increase in observations throughout the specified period relative to Bloomberg. This suggests that Refinitiv's ESG data is more representative of our Benchmark Index. Furthermore, the mean ESG scores range between 47.88 and 49.52 in our four-year period, with a substantially higher maximum score than Bloomberg of 92.97 in 2021.

Table 4.3: Descriptive Statistics: Sustainalytics ESG Scores

This table presents descriptive statistics of Sustainalytics' available ESG scores for firms in our sample. The observation period spans the years 2018 through 2021, with a monthly frequency, and the statistics are denoted as yearly averages. The table describes the total number of rated firms N(Firms), mean ESG scores (Mean), standard deviation (SD), minimum- (Min) and maximum (Max) values, and the 25^{th}, 50^{th}, and 75^{th} percentiles.

	2018	**2019**	**2020**	**2021**
N(Firms)	861.00	1148.00	1194.00	1282.00
Mean	44.22	52.74	69.61	70.71
SD	10.66	8.85	11.33	10.96
Min	7.18	27.97	28.19	27.86
25%	37.00	47.19	62.94	64.67
50%	46.50	53.00	70.59	71.37
75%	52.83	57.88	76.35	77.66
Max	66.00	86.35	99.88	99.90

The descriptive statistics of Sustainalytics are depicted in Table 4.3. Sustainalytics has the highest amount of firm observations for the entire period, where close to all the firms of our total sample size are covered by the end of 2021. Moreover, Sustainalytics' mean ESG scores are much higher at 70.71 in 2021 compared to both Bloomberg and Refinitiv. This is further substantiated by a maximum score of almost 100 for the same year.

4.6.1 Calculated Mean ESG Scores

Table 4.4: Descriptive Statistics: Calculated Mean ESG Scores

The following table presents descriptive statistics for the mean ESG scores based on the ESG scores from all rating providers. The purpose of this table is to give an overview of the calculated mean ESG scores across the three rating providers (i.e., Bloomberg, Refinitiv, and Sustainalytics) main attributes, the total number of observations N(Firms), mean ESG scores (Mean), standard deviation (SD), minimum- (Min) and maximum (Max) values, and the 25^{th}, 50^{th}, and 75^{th} percentiles for each year in our period respectively.

	2018	**2019**	**2020**	**2021**
N(Firms)	984.00	1178.00	1242.00	1307.00
Mean	44.13	49.56	57.63	59.34
SD	11.44	11.46	16.62	15.57
Min	1.10	1.25	5.21	10.39
25%	38.23	42.89	46.43	48.65
50%	45.96	50.85	57.88	59.61
75%	52.26	57.08	68.97	70.11
Max	73.12	86.36	99.88	99.90

Table 4.4 presents descriptive statistics for the calculated mean ESG score across our three rating providers. We identify a trend with increasing ESG scores across all the presented metrics in the period up to 2021, except for the standard deviation from 2020 to 2021. We also observe an increase in the number of covered firms. However, we can also observe an increase in the standard deviation over the entire period.

4.7 Portfolio Construction

To address our research questions and test our hypothesis, we generate three different portfolios: i) a *High* portfolio, which holds a long position in high ESG-rated firms, ii) a *Low* portfolio, which holds a long position in low ESG-rated firms, and iii) a *High-Low* portfolio[9] which represents the difference in returns between the high- and low portfolios.

[9]I.e., mimicking a portfolio that goes long in high ESG-rated firms and short in low ESG-rated firms.

Each portfolio is constructed by first defining 1-year periods[10], which is used for an annual rebalancing of the portfolio. The mean ESG score within each (yearly) period is attributed to each firm for each respective period. This normalizes any potential changes in ESG ratings for each firm within the annual period. The calculation steps are presented in Equation (4.3):

$$Mean\,ESG\,Score_{i,T} = \frac{ESG\,Score_{i,t_1} + ... + ESG\,Score_{i,t_{12}}}{12}, \tag{4.3}$$

where $Mean\,ESG\,Score_{i,T}$ is the calculated mean ESG score, t denotes the month, T denotes the entire time period, and i denotes the firm.

The mean ESG rating for each firm is then used to calculate the top (90^{th} and 75^{th}) and bottom (10^{th} and 25^{th}) percentiles within each period. These percentiles represent the restrictiveness or cutoffs when defining the high- and low ESG-rated firms for each portfolio, for that specific period[11]. The resulting portfolios are henceforth referred to as Decile and Quartile. The percentiles cutoffs of each portfolio were determined to examine whether the effect of ESG performance changed depending on how restrictive the definition of high- and low ESG-rated firms was.

Deciding whether to include the firm in our portfolio and whether it should be included in the high or low position is determined by assigning one of three numerical variables to each firm based on its ESG score for each period. This means that the high ESG-rated firms are assigned a 1, while the low ESG-rated firms are assigned a -1, where the numeric variable represents the denominator of the firm. The firms with ratings outside of the cutoffs are thus excluded[12] from the portfolio for that period and assigned a 0. Then we construct the high- and low portfolios, which contain only the firms in which we are invested in that period, where the assigned variables (1 and -1) represent whether the firm belongs in the high or low portfolio, respectively. Thus, the resulting high (low) portfolio reflects a portfolio with perceived high (low) sustainability performance, which allows us to determine the monthly abnormal returns per period for each respective portfolio, as well as the difference between the two. Each portfolio is then rebalanced annually,

[10] I.e., the first period ranges from January 2018 to December 2018 and so forth.

[11] E.g., the top (90^{th}) and bottom (10^{th}) cutoffs for the Decile is more restrictive, i.e., it requires firms to have a higher or lower ESG score (relative to the sample) to be included in the portfolio.

[12] See section Section 8.1.2 for an elaboration on addressing survivorship bias.

meaning that we allow firms to move in and out of the portfolio solely based on their ESG rating for each period. We generate the high-, low-, and high-low portfolios using both the Decile and Quartile.

Ultimately, each of our high-, low-, and high-low portfolios, constructed using both Decile and Quartile, respectively, includes four annual periods where portfolios are rebalanced. Thus, resulting in a total of 24 unique portfolios[13] across the 4-year time period, where each annual period is composed of 12 (monthly) observations for each firm.

4.8 Portfolio Return

When calculating portfolio return, we use the Total Return[14] (in U.S. dollars). We calculate both the equal- and value-weighted returns for each portfolio. The equal-weighted portfolio returns are simply calculated as the average monthly return of all firms included in the portfolio at time t.

For the value-weighted portfolio, the portfolio returns are a weighted average where the return of each firm is calculated based on the market capitalization of firm i, relative to the total sum of market capitalization for all included firms in the portfolio, at time t. The steps to calculate the value-weighted portfolio return are described in Equation (4.4):

$$w_{i,t} = \frac{Market\,Cap_{i,t}}{\sum_{i=1}^{N} Market\,Cap_{i,t}} \qquad r_{p,t} = \sum_{i=1}^{N} (w_{i,t} \times r_{i,t}), \tag{4.4}$$

where $w_{i,t}$ denotes the weight of firm i, $Market\,Cap_{i,t}$ denotes the market capitalization of firm i, $r_{p,t}$ denotes the value-weighted portfolio return r, and $r_{i,t}$ denotes the return of firm i, at time t.

Finally, in order to get the portfolio returns in excess of the risk-free rate, we subtract the risk-free rate[15] from the calculated high- and low portfolio returns respectively.

[13] We get 4 annually rebalanced portfolio compositions for each high-, low-, and long-short strategy (high-low), resulting in 12 unique portfolio compositions, multiplied by the 2 different thresholds (i.e., Decile and Quartile).

[14] See description of *Total Return* in Section 4.2.1.

[15] I.e., the one-month T-bill rate collected from French (2022) Data Library.

4.8.1 High-Low Difference Portfolio

Our generated high (low) portfolio consists only of the high (low) ESG-rated firms per period. Hence, by applying a zero-investment strategy, the low portfolio return is then *subtracted* from the high portfolio return. Accordingly, by considering their differences in return, similar to the approach used by Kempf and Osthoff (2007) and Gibson Brandon et al. (2021), our high-low portfolio return is designed to capture and reflect some of the influence of ESG performance on stock returns. The calculation steps to calculate the high-low portfolio return are shown in Equation (4.5):

$$r_{High-Low,t} = R_{High,t} - R_{Low,t}, \tag{4.5}$$

where $r_{High-Low,t}$ denotes high-low portfolio return, $R_{High,t}$ denotes the high portfolio excess return, and $R_{Low,t}$ denotes the low portfolio excess return, at time t. In contrast to the high- and low portfolio returns, we do not deduct the risk-free rate from the high-low portfolio, as a zero-investment strategy has a net-zero alternative cost. Additionally, as described in Section 4.8, we also calculate both the equal-weighted and value-weighted high-low portfolio return, where the latter aim to adjust for any potential firm-specific effects related to market capitalization.

4.9 Fama-French Factors

The Fama-French factors were downloaded from Kenneth R. French's data library for emerging markets. All returns include dividends and capital gains and are stated in U.S. dollars, but they are not continually compounded (French, 2022). Instead, the U.S. one-month treasury bill rate is used by Kenneth R. French Library by default as a risk-free return, which we found best suited for the purpose of our thesis. Conveniently, the countries included in the Kenneth R. French data library are almost identical to our benchmark index, the MSCI Emerging Markets Large & Mid Cap index[16].

[16]See section 4.3 for more details.

4.10 Concerns about the Data

The following section discusses how we managed different concerns about the data used in our analysis. Therefore, in our interpretations, we have been conscious of the potential effects of the data limitations.

4.10.1 Choice of Determinant for Financial Performance

We could have used numerous financial metrics to capture different aspects of financial performance. For instance, a significant portion of corresponding literature uses accounting terms like Return on Equity (ROE) and Return on Assets (ROA) to measure financial performance (Duque-Grisales & Aguilera-Caracuel, 2021). Accounting metrics are generally good indicators of a firm's profitability and are easily accessible through financial reports. Thus, one could argue that these metrics would serve as a more suitable alternative and are more representative of the literature compared to the financial variables we applied, i.e., *Total Return* and *Market Capitalization*. However, equity returns and company market valuations (market values) have merits beyond accounting variables (book values) because they reflect the sentiment of financial markets. Market-based metrics should therefore be more indicative of future firm performance by reflecting shareholders' expectations. Consequently, market-based financial metrics are preferable in our context due to the methodology applied and the comparison with the benchmark index regarding abnormal returns.

4.10.2 Variations in Methodology Between Rating Providers

There is an absence of common standards between rating providers on how to evaluate a firm's ESG performance. This includes differences in the frequency of when the providers release updated ESG ratings, reporting standards, and the number of firms covered by providers in emerging markets. Consequently, we are faced with some concerns regarding the validity and degree of subjectivity among our rating providers. Also, some of the data providers may have changed their respective rating methodologies during our sample period[17], which would generate additional biases. However, by calculating the mean ESG score, we assume some of the abovementioned differences were mitigated.

[17]Assuming they do not update historical scoring data.

4.10.3 Deviating Actual and Expected Index Performance

One challenge associated with the expected benchmark index (MSCI Emerging Markets Index) is the fact that it gives no information regarding historical constituents, i.e., we only have a snapshot of the index constituents. As indexes are generally weighted differently based on their respective market capitalization weightings, the composition of the actual benchmark index will therefore shift frequently, consistent with share price movements. Meanwhile, the composition of our expected benchmark index, which only presents a snapshot of the index cross-section, remains unchanged (as of the last update, September 1, 2022). This also implies that our sample of 1327 constituents is likely to change noticeably over time. Consequently, if this exercise was to be replicated in the future, deviations from our results ought to be expected[18].

Furthermore, we use Bloomberg's Emerging Markets Index as a proxy for the MSCI Emerging Markets Index, which is designed to replicate overall development in emerging markets. However, there are 59[19] fewer constituents included in the Bloomberg index. As a result, it is not unlikely that there exist some variations in comparison to MSCI's Emerging Markets Index. Nevertheless, despite these differences, we are compelled to make an underlying assumption that Bloomberg's index is representative of the link between ESG and CFP in emerging markets. We are also convinced that this assumption holds due to the large number of constituents in our sample index.

4.10.4 Concerns regarding the Sample Selection

There are some countries included in the Fama-French data[20] that are not included in MSCI's Emerging Markets Index[21]. Hence, there may be some undesirable regional deviations caused by the memberships of Russia, Qatar, and Argentina in the Fama-French data.

[18]Where the extent of deviation is dependent on the corresponding change in index composition.

[19]MSCI's EM index constituents (1386) - Bloomberg's EM index constituents (1327).

[20]Per September 2022, the Fama-French emerging markets countries include: 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, and United Arab Emirates (French, 2022).

[21]For comparison with countries in the MSCI Emerging Markets Index, see Section 4.3.

4.10.5 Deviations in Rating Availability

Additionally, the number of ESG-rated firms (i.e., the number of observations) differs widely per year and between the three rating providers. As we rely upon the ESG ratings to construct our portfolio, the number of missing observations across the three providers could then have posed a problem if we were to consider each provider's scores separately. For instance, out of the total 1327 constituent firms in our sample, Bloomberg only provides ESG scores for 187 firms for the year 2021, while Refinitiv and Sustainalytics provide ESG ratings for a total of 1178 and 1282 firms, respectively[22]. However, this issue is partially mitigated by the fact that we use the mean of ESG scores across all available ratings for each firm in our sample rather than considering ESG scores from each rating provider independently. Consequently, compared to Refinitiv and Sustainalytics, which have far more observations, Bloomberg's methodology is given less significance in our analysis. Furthermore, this implies that our data is more exposed to subjectivity from Refinitiv and Sustainalytics. Ultimately, we end up with an ESG score that better represents the overall consensus of our three providers[23].

A central criterion of our data preparation is that we include firms with only one available score (i.e., assign a firm's ESG score based on data from solely one provider). For example, we could have required that all three providers had available ESG scores for a company to be included in our constructed portfolio as an alternative to our preferred approach. Ideally, this would ensure higher confidence in our data compared to the instances where only one observation is available, as the observation would reflect the average of three ESG scores. However, in reality, our requirement of only one provider was conducted to ensure that companies were included in the respective portfolio rather than to be excluded and treated as missing values (NA). Hence, a stricter requirement of more than one provider would lead to more missing data, thus making our analysis less robust. Ultimately, it is a trade-off between more observations (by utilizing all available data through our one-provider criterion) on one side and more bias (the provider with the most data points is more frequently included) on the other. In the context of our overarching research question, we are convinced that the latter is less important because we value more observations to reflect the diversity of firm ESG performance in emerging markets.

[22] See Section 4.6.1.

[23] Consistent with the discussion in Section 2.4.1.

Moreover, we find it a realistic assumption that most investors will view having only one rating provider as sufficient when making investment decisions based on ESG performance.

5 Methodology

This chapter aims to describe and explain the methodology choices taken to conduct our analysis. Initially, we present the Fama-French model framework and the four different risk-factor regression models used in our analysis. A brief discussion on the choice of models follows before we state the underlying model assumptions. Finally, we present and discuss the tests performed on each model to ensure a reliable interpretation of the regression results.

5.1 Factor Regressions

To capture the effect of ESG ratings on excess returns, we run factor regressions on our high-, low-, and high-low portfolios using Ordinary Least Squares (OLS) regression. The dependent variables are the calculated equal- and value-weighted portfolio returns in excess of the risk-free rate[24] for each of our high-, low-, and high-low portfolios.

5.2 Fama-French Model Specifications

The Fama-French model framework, which expands on the Capital Asset Pricing Model (CAPM), is widely recognized and used frequently by academia and investors to analyze market performance. The Fama-French model framework consists of several factors which aim to explain the variation in excess return (Womack & Zhang, 2003). The objective of the Fama-french factors is thus to capture all variations in the market. In addition, as the Fama-French framework is widely recognized and used by both academia and investors, we can compare our results to the existing literature.

When interpreting the results, an intercept (i.e., alpha) of zero implies that the long-short strategy creates no abnormal returns. For instance, when considering the high-low portfolio, an intercept of 1 (-1) means that the high (low) ESG-rated firms outperform, in terms of monthly returns, the low (high) ESG-rated firms. Moreover, the statistical significance

[24]As described in Section 4.8.1, the Risk-free rate is only subtracted from the high- and low portfolios.

of estimated coefficients implies a difference, between the high- and low portfolios, in exposure to the specific risk factor. Consequently, when considering the differences in a zero-investment portfolio, the estimated coefficients may have less explanatory power, and we will thus focus our interpretation on the coefficients denomination.

5.2.1 Fama-French Three-Factor Model

The first iteration of the Fama-French model framework is the Fama-French Three-Factor model, published by Fama and French (1992). The three-factor model expands on the framework laid forth by the CAPM framework. The market factor (MktRf) denotes the total market portfolio return in excess of the risk-free rate, while the firm-specific factors "Small Minus Big" (SMB) and "High Minus Low" (HML) represent the risk premium related to firm size and value, respectively. The formula used for the Fama-French Three-Factor Model regression is shown in Equation (5.1):

$$r_{p,t} = \alpha + \beta_{MktRf}(Mkt_t - r_{f,t}) + \beta_{SMB}(SMB_t) + \beta_{HML}(HML_t) + \epsilon_t, \tag{5.1}$$

where $r_{p,t}$ is the portfolio return in excess of the risk-free rate[25], α is the intercept or abnormal return (i.e., alpha), β_{MktRf} denotes the exposure to the market factor and $Mkt_t - r_{f,t}$ is the market premium, β_{SMB} denotes the exposure to the size factor and SMB_t is the size premium, β_{HML} denotes the exposure to the value factor and HML_t is the value premium, and ϵ_t denotes the error term, at time t.

5.2.2 Carhart Four-Factor Model

The Carhart Four-Factor Model expands on the Fama-French Three-Factor Model and adds the momentum factor, denoted as "Winners Minus Losers" (WML). This factor seeks to capture the persistence in firm performance and represents the risk premium related to momentum exposure (Carhart, 1997). The formula used for the Carhart Four-Factor regression is shown in Equation (5.2):

[25]The risk-free rate used is the U.S. one month T-bill rate, collected from Kenneth R. French (2022) Data Library.

$$r_{p,t} = \alpha + \beta_{MktRf}(Mkt_t - r_{f,t}) + \beta_{SMB}(SMB_t) + \beta_{HML}(HML_t) + \beta_{WML}(WML_t) + \epsilon_t, \quad (5.2)$$

where β_{WML} is the exposure to the momentum factor, and WML_t is the momentum premium at time t.

5.2.3 Fama-French Five-Factor Model

In 2015, Fama and French published a research paper that revised the Fama-French three-factor model by expanding the model to include two additional factors, "Robust Minus Weak" (RMW) and "Conservative Minus Aggressive" (CMA) (Fama & French, 2015). The RMW factor denotes the difference in returns between firms with robust and weak profitability. The CMA factor indicates the difference in returns between firms that invest conservatively and those that invest aggressively. According to Fama and French's research, the Five-Factor Model performed better than the Three-Factor Model in terms of capturing size, value, profitability, and investment patterns in average stock return (Fama & French, 2015). The formula used for the Fama-French five-factor regression is shown in Equation (5.3):

$$\begin{aligned} r_{p,t} = \alpha + \beta_{MktRf}(Mkt_t - r_{f,t}) + \beta_{SMB}(SMB_t) + \beta_{HML}(HML_t) \\ + \beta_{RMW}(RMW_t) + \beta_{CMA}(CMA_t) + \epsilon_t, \end{aligned} \quad (5.3)$$

where β_{RMW} is the exposure to the profitability factor and RMW_t is the profitability premium, β_{CMA} is the exposure to the investment factor, and CMA_t is the investment premium, at time t.

5.2.4 Fama-French Five-Factor Model with Momentum

The Fama-French Five-Factor Model with momentum (WML) factor expands on the original Five-Factor Model described in subsection 5.2.3 and adds the momentum factor, which is described in subsection 5.2.2. The formula used for the Fama-French Five-Factor with momentum regression is shown in Equation (5.4):

$$r_{p,t} = \alpha + \beta_{MktRf}(Mkt_t - r_{f,t}) + \beta_{SMB}(SMB_t) + \beta_{HML}(HML_t) + \beta_{RMW}(RMW_t) + \beta_{CMA}(CMA_t) + \beta_{WML}(WML_t) + \epsilon_t. \tag{5.4}$$

5.3 Model Testing

When conducting time-series analysis and regression, we need to assert that we are dealing with stationary processes. Therefore, any potential non-stationary processes need to be transformed (to a stationary process) before being used in linear regression. An augmented Dickey-Fuller test[26] was used to check for stationarity (Wooldridge, 2020). The number of lags used is determined automatically by choosing the number of lags that yields the lowest Akaike's information criterion (AIC). The results of the augmented Dickey-Fuller test are presented in Table A1.1 and Table A1.2 for Decile and Quartile, respectively, which shows that all processes are stationary at the 5% level.

5.3.1 OLS Assumptions

To avoid spurious regression results and thus to assure a valid interpretation of the results, the underlying data needs to satisfy certain assumptions: i) linear parameters, ii) no perfect collinearity, iii) zero conditional mean, iv) homoskedasticity, and v) no autocorrelation (Wooldridge, 2020). The estimators are unbiased if assumptions i)-iii) are satisfied, and if assumptions i-v) are satisfied, then the estimators are denoted as the Best Linear Unbiased Estimators (BLUE).

5.3.2 Multicollinearity

The Variance Inflation Factor (VIF) test was conducted to identify multicollinearity, and the results are presented in Table A1.7. VIF, which exceeds 10, generally implies problematic multicollinearity, while values above 4 indicate correlation that warrants further investigation (Wooldridge, 2020). However, our VIF results show that all our values are below 4. Hence we do not have a problem with multicollinearity.

[26]The augmented version of the Dickey-Fuller test is used to account for any possible autocorrelation.

5.3.3 Heteroskedasticity and Autocorrelation

The Breusch-Pagan test was conducted to formally test for heteroskedasticity, while the Breusch-Godfrey test was used to identify autocorrelation. The Breusch-Pagan test results are presented in Table A1.5 for Decile and Table A1.6 for Quartile, while the Breusch-Godfrey test results are presented in A1.3 for Decile and Table A1.4 for Quartile. We find no immediate evidence of heteroskedasticity from the results of the Breusch-Pagan tests.

However, we identify one incident (i.e., the Fama-French Five-Factor Model with Momentum for Quartile[27]) of autocorrelation from the Breusch-Godfrey test for this model. Nonetheless, given that this one-time occurrence, we are not overly concerned about its implications for our analysis.

Furthermore, we used linear regression diagnostic plots[28] to inspect whether our regression models satisfied the assumptions for OLS regression. The plots are presented in A2.1, A2.2, and A2.3 in the Appendix.

We identify minor deviations from normality and heteroskedasticity. However, we find no dramatic tendencies for either. Although the trends for heteroskedasticity were more pronounced, we ran the regressions with the heteroskedasticity robust standard errors (HC1) test. Moreover, due to the COVID-19 pandemic, the time period analyzed is characterized by a high degree of financial market volatility. Thus, some deviations from both normality and heteroskedasticity are likely to be expected.

[27] See Table A1.4 in Appendix.
[28] See Appendix A2 for further description.

6 Analysis

In this section, we present the results of our analysis. As such, we aim to determine if our high (low) portfolio, going long (short) in high (low) ESG-rated firms, results in significant abnormal returns by comparing these portfolios through a high-low (i.e., a long-short) difference portfolio. This is done for Decile, Quartile, and for equal- and value-weighted portfolio returns, respectively. Initially, we present descriptive statistics on portfolio performance and industry sector composition for each portfolio. Then, we present the regression results of our constructed portfolios using each respective model specification, as outlined in Section 5.

6.1 Descriptive Results

The following section presents the descriptive results of our analysis, where we present an overview of the attributes of our portfolios.

6.1.1 Portfolio Statistics

Table 6.1: Descriptive Portfolio Results

This table presents descriptive results for each equal-weighted High-, Low-, and High-Low (i.e., H-L) portfolios, decomposed into the Decile and Quartile cutoffs. The sample includes 48 monthly time series observations in the period January 2018 to 2021, which are aggregated by year. For both cutoffs and each respective portfolio, we present the following: *N(Firms)* denotes the number of included firms in each High- and Low- portfolio (here H&L represents both High and Low, as they include an equal number of firms per period), as well as the difference portfolio High-Low; *Avg. Market Cap.* denotes the average market capitalization of the included firms, denoted in billion USD; *Avg. Returns (%)* denote the annualized (monthly) average returns (in percentage); and *Avg. Sharpe Ratio* denotes the annualized (monthly) average Sharpe Ratio of each portfolio. To highlight the yearly differences within the sample period, we apply a color gradient to emphasize the performance for the given year relative to the overall period, where darker gradients denote higher values. The color mapping is applied to the high- and low portfolios, where the former is colored in shades of blue and the latter in shades of green.

	N(Firms)		Avg. Market Cap.			Avg. Returns (%)			Avg. Sharpe Ratio		
	H & L	**H-L**	**High**	**Low**	**H-L**	**High**	**Low**	**H-L**	**High**	**Low**	**H-L**
Decile											
2018	99	198	29.6	8.2	21.4	0.16	-3.74	3.90	-0.08	-0.44	0.13
2019	118	236	14.2	9.2	5.0	5.59	5.00	0.59	0.50	0.38	-0.02
2020	125	250	19.1	9.8	9.3	9.81	16.73	-6.92	0.73	1.39	-0.31
2021	131	262	25.1	13.1	12.0	16.51	12.43	4.09	0.98	0.79	0.06
Quartile											
2018	246	492	26.1	12.2	14.0	-0.66	-3.03	2.38	-0.14	-0.38	0.06
2019	295	590	17.7	10.2	7.5	5.13	4.79	0.34	0.52	0.35	-0.03
2020	311	622	15.3	16.6	-1.3	6.72	13.11	-6.39	0.55	0.97	-0.24
2021	327	654	20.8	19.8	0.9	10.46	8.48	1.98	0.78	0.59	0.05

Table 6.1 presents annual statistics of the generated high-, low- and high-low portfolios for each respective cutoff. When examining the size of each portfolio, we see that the number of firms included per year is identical for each respective high- and low portfolio[29]. Moreover, we observe a consistent increase in the included number of firms for all portfolios, irrespective of cutoffs. This growth can generally be credited to either i) the number of firms that satisfy the strictest cutoff requirement, with either good or bad ESG practices, is more common or ii) ESG providers[30] have expanded their coverage of firms. Moreover, we observe that the Quartile portfolio includes more than double the number of firms for all periods, relative to its Decile counterpart, due to the restrictiveness criterion described in Section 4.7.

When considering the average market capitalization for each year in figure 6.1, we observe

[29] As described in Section 4.7.
[30] I.e., Bloomberg, Refinitiv, and Sustainalytics.

that the high portfolio has significantly greater market capitalization for all years for Decile, apart from the year 2019 for Quartile. Moreover, the positive difference in average market capitalization between the high- and low portfolios is especially pronounced for the Decile cutoff, which implies that the top Decile portfolio of high ESG-rated firms, on average, is firms with higher market value. Conveniently, the positive difference in average market capitalization turned negative for the Quartile cutoff in 2020, when the COVID-19 market turmoil struck the financial markets. However, the positive difference in 2020 in average market capitalization persisted for the Decile cutoff, suggesting that some high ESG-rated firms were less affected (in terms of market value) by the financial instability relative to low ESG-rated firms in the Quartile cutoff.

When examining the average returns for each period, we observe that the high portfolio consistently yields higher average returns for all periods (except 2020) compared to the low portfolio. As such, high ESG-rated firms outperformed (in terms of average annualized monthly returns) low ESG-rated firms for the years before and after the COVID-19 crisis took place, regardless of cutoffs.

When considering the Sharpe Ratio for each period, we notice that the risk-adjusted returns were better for the high portfolio relative to the low portfolio for both cutoffs (except for the year 2020). This implies that the high portfolio has a higher return for a given level of risk for the years 2018, 2019, and 2021. However, the highest Sharpe Ratio was observed for the Decile low portfolio in 2020. This is most likely driven by higher returns, as reflected in unusually higher average returns (relative to our sample period) compared to each level of risk.

6.1.2 Portfolio Returns

An introductory part of our descriptive analysis is to plot the cumulative returns of our constructed portfolios relative to the market premium for emerging markets.

Figure 6.1: Cumulative Monthly Returns of the High-, Low-, and High-Low portfolios
This figure presents the cumulative monthly returns (null indexed, i.e., starting at 0%) of our constructed zero-investment ESG portfolios from January 2018 to December 2021. The figure is decomposed into three subplots, all of which contain the market risk premium (denoted MktRF) and describes the cumulative returns throughout our period for the following sets of portfolios; *i) High-Low*, presents our constructed equal- and value-weighted high-low Decile and Quartile; *ii) Equal-Weighted High and Low*, presents our constructed equal-weighted high- and low portfolios; and *iii) Value-Weighted High and Low*, presents our constructed value-weighted high- and low portfolios.

Figure 6.1 displays the cumulative monthly returns (null indexed) of our three constructed portfolios, high-, low- and high-low portfolios, and the market risk premium (MktRF), decomposed into the equal- and value-weighted Decile and Quartile. When considering the high-low portfolio presented in the first subplot *i) High-Low*, the equal-weighted Decile outperforms (i.e., yields the highest returns) its value-weighted counterpart, whereas the value-weighted Quartile marginally outperforms the equal-weighted Quartile. Both Decile weightings outperformed the Fama-French market risk premium from approximately July 2018 until July 2020, after the disruptive effects of COVID-19 on the stock market began

to wear off.

When examining the equal-weighted portfolios in the second subplot *ii) Equal-Weighted High and Low*, we observe a clear distinction where both the high- and low Decile portfolios significantly outperform their Quartile counterparts for the entire period, with a few exceptions in the period March-April 2020, which is likely explained by the COVID-19-induced market volatility.

We can identify similar movements for the value-weighted portfolios in the third subplot *iii) Value-Weighted High and Low*, albeit there is less volatility which suggests that the value-weighted portfolios are more robust to market turmoil than the equal-weighted. When value-weighted, Decile outperforms Quartile on average (although not consistently), particularly after the financial turmoil propelled by COVID-19. Relative to equal-weighted, there is a greater divergence in cumulative returns after April 2020 (when the COVID-19-induced stock crash bottomed out), and positive market sentiment followed. This especially holds for the value-weighted high- and low Decile, which proved more robust under the market upsurge relative to their Quartile counterparts.

6.1.3 Portfolio Sector Composition

Analyzing the sector distribution helps to provide insight into the industry sector exposure of our portfolio and thus be useful to consider diversification, identify investment opportunities, and market trends or risks.

To get an overview of the sector distribution among our high-, low-, and high-low portfolios, we define and classify securities by industries using the Global Industry Classification Sector (GICS) framework (MSCI, 2020). Figure 6.2 presents a sector composition breakdown, highlighting industrial sector exposure for each respective portfolio decomposed into Decile and Quartile cutoffs, where *i) Decile Cutoff* presents the Decile distribution for the high-low portfolio, decomposed in the respective high- and low portfolios; *ii) Quartile Cutoff* presents the Quartile distribution for the high-low portfolio, decomposed in the respective high- and low portfolios; *iii) High Portfolio* shows the high portfolio sector composition, decomposed into Decile and Quartile cutoffs; and *iv) Low Portfolio* shows the low sector composition, decomposed into Decile and Quartile cutoffs.

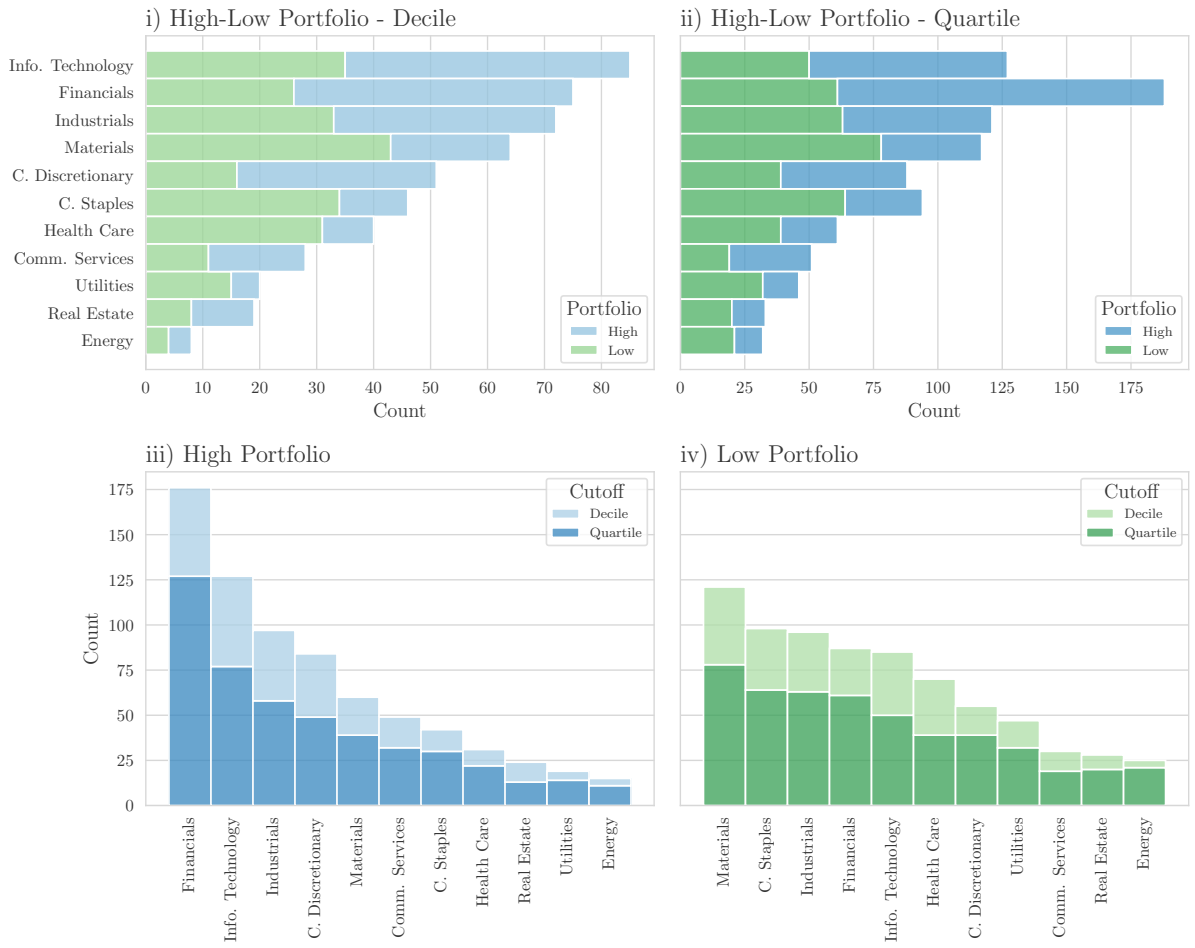

Figure 6.2: Global Industry Classification Sector Composition
This figure illustrates our portfolio compositions using the Global Industry Classification Sectors (GICS) for *i) High-Low Portfolio - Decile* and *ii) High-Low Portfolio - Quartile* cutoffs for the high-low portfolio, decomposed into the *iii) High Portfolio* and *iv) Low Portfolio*. The figure provides a breakdown of portfolio industry exposure by grouping the firms in our respective portfolios by industrial sectors. The first two subplots display the stacked high-low portfolio distributions for the *Decile* and *Quartile* cutoffs. In contrast, the subsequent two portfolios display the stacked Decile and Quartile cutoffs for the *high-* and *low* portfolios. All subplots depict the most frequently occurring sectors and are sorted in descending order (except the Quartile cutoff, based on the order of the Decile cutoff). In no particular order, the GICS consists of the following 11 sectors: Financial, Industrials, Information Technology, Materials, Consumer Staples, Consumer Discretionary, Health Care, Utilities, Communication Services, Energy, and Real Estate.

As presented in Figure 6.2, the *i) High-Low Portfolio - Decile* and *ii) High-Low Portfolio - Quartile* subplots denote the overall number of firms included in the Decile and Quartile cutoffs, respectively. As such, we notice that an increased number of firms when considering *ii) High-Low Portfolio - Quartile*, results in a more unevenly distributed exposure to different sectors. Moreover, when examining the first two subplots, we find certain similarities between the stacked high- and low portfolios. For instance, the top three most frequently occurring sectors across both cutoffs are Information Technology, Financials,

and Industrials. However, in contrast to Decile, the Financials sector assumes the top spot over Information Technology when considering Quartile. Finally, the Consumer Discretionary and Consumer Staples sectors are typically the 5^{th} and 6^{th} most frequently occurring sectors across both Decile and Quartile cutoffs. This could be a reflection of the expanding opportunities in emerging markets amid the rapidly growing middle class with greater disposable income in the consumer sectors (Grohmann, 2018).

Furthermore, comparing figures *iii) High Portfolio* and *iv) Low Portfolio* brings forth differences in sector composition of high ESG-rated firms and low ESG-rated firms sorted in descending order by total occurrences in Decile and Quartile. We observe a distinct overweight of firms in the Financials sector among the high portfolio, whereas the exposure to the Financials sector is less pronounced in the low portfolio. In addition, Information Technology constitutes the second most frequently occurring sector among the high portfolio, while it is only the 5^{th} most common sector for the low portfolio. Also, the overall distribution of the high portfolio is more left-skewed (i.e., biased through a higher concentration of firms within a small selection of sectors), while the low portfolio is more evenly distributed.

Interestingly, a common denominator for the two portfolios is that the Energy sector is the most underrepresented sector across both the high- and low portfolios. Therefore, due to relatively lower exposure to the Energy sector, it might infer weaker external validity. Overall, the most notable differences between the sector composition of the high- and low portfolio seem to be that high ESG-rated firms have an increased orientation toward the Financials and Information Technology sector relative to low ESG-rated firms.

6.2 Regression Results

The following part presents our results from the regression analysis. We only interpret results deemed to be significant, i.e., statistical significance at the 5% level or below. Hence, we do not elaborate or interpret any results deemed statistically insignificant, i.e., above the 5% level. For consistency, all regressions are conducted for both Decile and Quartile, but for Quartile, we limit the focus of our interpretation to the observed differences relative to Decile.

6.2.1 Fama-French Three-Factor Model

Table 6.2: Fama-French Three-Factor Model Decile Regression Results

This table presents the Fama-French Three-Factor model regression results for the Decile portfolio. The intercept coefficients denote the abnormal monthly return (in percentage) for the equal- and value-weighted high-, low-, and high-low portfolios. The monthly risk-free rate (one-month U.S. treasury bill) is deducted from both the high- and low portfolios, while the *difference* portfolio (high-low) is constructed by subtracting the low portfolio returns from the returns of the high portfolio. The dependent variable is the monthly return ($r_{p,t}$) for each respective portfolio, and the independent variables represent the risk factors MktRF, SMB, and HML. Heteroskedasticity Robust (HC1) Standard Errors are represented in parenthesis below each respective coefficient. The sample includes 48 monthly time series observations in the period January 2018 to 2021.

	Dependent Variable: Portfolio Return					
	High		**Low**		**High-Low**	
	EW	VW	EW	VW	EW	VW
Intercept	1.692***	0.725***	1.911***	1.047***	-0.219	-0.322
	(0.325)	(0.282)	(0.466)	(0.208)	(0.391)	(0.314)
MktRF	1.068***	0.671***	0.881***	0.322***	0.187**	0.348***
	(0.059)	(0.060)	(0.100)	(0.043)	(0.088)	(0.050)
SMB	0.632***	-0.387**	0.614**	0.041	0.018	-0.428**
	(0.219)	(0.195)	(0.279)	(0.113)	(0.244)	(0.190)
HML	0.328**	0.247***	-0.548***	-0.240***	0.876***	0.486***
	(0.136)	(0.087)	(0.183)	(0.079)	(0.150)	(0.097)
Observations	48	48	48	48	48	48
R^2	0.881	0.799	0.691	0.586	0.538	0.569
Adjusted R^2	0.873	0.785	0.670	0.558	0.506	0.539
Note:					*p<0.1; **p<0.05; ***p<0.01	

When considering Decile, the regression output for the Fama-French Three-Factor model is presented in Table 6.2 and describes positive abnormal returns (i.e., the intercept or α) in the equal- and value-weighted portfolios for both high- and low portfolios, which are significant at the 1% level. As such, the positive intercepts of the equal-weighted high (low) portfolio indicate that high (low) ESG-rated firms generate monthly abnormal returns of 1.692% (1.911%) in excess of the risk-free rate. Meanwhile, the monthly abnormal return is halved when considering its value-weighted counterparts. However, we observe no significance at any level for either the equal- or value-weighted high-low portfolio, thus suggesting that the difference in ESG performance is not able to explain the variation in returns using this model specification.

We observe a positive and significant MktRF factor (i.e., the market risk premium) at

the 1% level for all portfolios and weightings[31]. Furthermore, the market risk premium is highest for the high portfolio regardless of weighting, suggesting that the high ESG-rated firms generated a greater market risk-adjusted return relative to low ESG-rated firms.

The SMB factor (i.e., the size premium) varies in magnitude and significance for all portfolios and weightings. The size premium is positive and significant for the equal-weighted high- and low portfolios, which implies a bias toward small-cap stocks. However, we observe a negative and significant size premium for the value-weighted high- and high-low portfolios, suggesting that the high ESG-rated firms generally have a greater market capitalization.

The HML factor (i.e., the value premium) is significant for all portfolios and weightings but only positive for the high- and high-low portfolios. As such, the positive (negative) value premium indicates an increased orientation towards value (growth) firms among the high (low) ESG-rated firms in our portfolio.

[31]Except the equal-weighted high-low portfolio, which is significant at the 5% level.

Table 6.3: Fama-French Three-Factor Model Quartile Regression Results

This table presents the Fama-French Three-Factor model regression results for the Quartile portfolio. The intercept coefficients denote the abnormal monthly return (in percentage) for the equal- and value-weighted high-, low-, and high-low portfolios. The monthly risk-free rate (one-month U.S. treasury bill) is deducted from both the high- and low portfolios, while the *difference* portfolio (high-low) is constructed by subtracting the low portfolio returns from the returns of the high portfolio. The dependent variable is the monthly return ($r_{p,t}$) for each respective portfolio, and the independent variables represent the risk factors MktRF, SMB, and HML. Heteroskedasticity Robust (HC1) Standard Errors are represented in parenthesis below each respective coefficient. The sample includes 48 monthly time series observations in the period January 2018 to 2021.

	Dependent Variable: Portfolio Return					
	High		**Low**		**High-Low**	
	EW	VW	EW	VW	EW	VW
Intercept	0.947***	0.433***	1.324***	0.685***	-0.378	-0.252
	(0.203)	(0.166)	(0.322)	(0.154)	(0.291)	(0.222)
MktRF	1.046***	0.558***	0.949***	0.399***	0.098	0.158***
	(0.034)	(0.041)	(0.066)	(0.031)	(0.072)	(0.043)
SMB	0.349***	-0.348***	0.312*	0.120	0.037	-0.469***
	(0.131)	(0.128)	(0.176)	(0.083)	(0.152)	(0.143)
HML	0.376***	0.048	-0.354***	-0.149**	0.730***	0.197***
	(0.080)	(0.050)	(0.121)	(0.063)	(0.116)	(0.075)
Observations	48	48	48	48	48	48
R^2	0.947	0.883	0.842	0.796	0.557	0.406
Adjusted R^2	0.944	0.875	0.832	0.783	0.527	0.365
Note:					*p<0.1; **p<0.05; ***p<0.01	

When considering Quartile, the regression results using the Fama-French Three-Factor model specification are shown in Table 6.3. By loosening the ESG score criterion of a high (low) ESG-rated firm[32], we observe a tendency of reduced magnitudes for the majority of factor coefficients for all portfolios, regardless of weightings.

Regarding the market risk premium, all factor coefficients remain significant except for the equal-weighted high-low portfolio. In contrast to Decile, we do not find a significant difference in the market risk premium for the equal-weighted high-low portfolio.

The observed significance and magnitude of the size premium remain similar to that of its Decile counterparts for the high- and high-low portfolios, regardless of weighting. However, in contrast to the Decile results, we find no statistical significance for either the equal- or

[32]I.e., changing the restrictiveness and thus increasing the number of firms included in the portfolio at any given time, by going from the Decile (Quartile) 10^{th} (90^{th}) to the 25^{th} (75^{th}) percentile cutoffs.

value-weighted low portfolio[33].

Finally, the value premium appears to be consistent across portfolios, with the exception of the value-weighted high portfolio, which has a statistically insignificant value premium when examining Quartile.

6.2.2 Carhart Four-Factor Model

Table 6.4: Carhart Four-Factor Model Decile Regression Results

This table presents the Carhart Four-Factor model regression results for the Decile portfolios. The intercept coefficients denote the abnormal monthly return (in percentage) for the equal- and value-weighted high-, low-, and high-low portfolios. The monthly risk-free rate (one-month U.S. treasury bill) is deducted from both the high- and low portfolios, while the *difference* portfolio (high-low) is constructed by subtracting the low portfolio returns from the returns of the high portfolio. The dependent variable is the monthly return ($r_{p,t}$) for each respective portfolio, and the independent variables represent the risk factors MktRF, SMB, HML, and WML. Heteroskedasticity Robust (HC1) Standard Errors are represented in parenthesis below each respective coefficient. The sample includes 48 monthly time series observations in the period January 2018 to 2021.

	Dependent Variable: Portfolio Return					
	High		**Low**		**High-Low**	
	EW	VW	EW	VW	EW	VW
Intercept	1.436***	0.414	1.455**	0.813***	-0.019	-0.398
	(0.440)	(0.334)	(0.622)	(0.266)	(0.436)	(0.372)
MktRF	1.090***	0.697***	0.920***	0.342***	0.170*	0.355***
	(0.054)	(0.061)	(0.097)	(0.043)	(0.094)	(0.055)
SMB	0.592***	-0.436**	0.543*	0.004	0.049	-0.440**
	(0.216)	(0.188)	(0.279)	(0.116)	(0.247)	(0.197)
HML	0.463*	0.411***	-0.307	-0.115	0.770***	0.527***
	(0.239)	(0.115)	(0.313)	(0.124)	(0.180)	(0.141)
WML	0.221	0.269**	0.394	0.203	-0.173	0.066
	(0.213)	(0.135)	(0.317)	(0.131)	(0.207)	(0.131)
Observations	48	48	48	48	48	48
R^2	0.886	0.814	0.709	0.617	0.546	0.570
Adjusted R^2	0.875	0.796	0.682	0.581	0.504	0.530
Note:					*p<0.1; **p<0.05; ***p<0.01	

When examining the Decile regression output for the Carhart Four-Factor model, as shown in Table 6.4, we find results that are largely consistent with those obtained when using the Fama-French Three-Factor model specification, even when controlling for the WML factor (i.e., the momentum premium). In addition, we observe positive abnormal returns

[33]Although the equal-weighted low portfolio is significant at the 10% level.

for the equal-weighted high- and low portfolios at the 5% level (or below). Moreover, we observe a less pronounced difference in abnormal returns between the equal-weighted high- and low portfolios[34], relative to the results using the Three-Factor model specification. Meanwhile, only the value-weighted low portfolio is significant at the 1% level. As such, the positive intercepts of the equal-weighted high (low) portfolio imply that high (low) ESG-rated firms generate monthly abnormal returns of 1.436% (1.455%) in excess of the risk-free rate. Finally, we observe no significance at any level for either the equal- or value-weighted high-low portfolio, which suggests that the difference in ESG performance does not explain the variation in returns using this model specification.

The MktRF factor is positive and significant at the 1% level for all portfolios and weightings[35]. Hence, in resemblance to the Decile results when applying the Three-Factor model, the market risk premium is greatest for the high portfolio irrespective of weightings, implying that the high ESG-rated firms generated a greater market risk-adjusted return compared to low ESG-rated firms.

When using the Four-Factor model specification for Decile, the SMB factor consistently displays a significant size premium for the high portfolio, similar to what was observed when applying the Three-Factor model. In terms of the equal-weighted high portfolio, the positive and significant size premium implies a greater fraction of small capitalization firms among the high ESG-rated firms. Meanwhile, its value-weighted counterpart denotes a negative and significant size premium, suggesting a greater orientation toward large capitalization firms when we adjust portfolio return for firm size. While also positive, none of the size premiums for any weightings of the low portfolio are significant. Conversely, we observe a negative and significant size premium in the value-weighted high-low portfolio, implying that the high ESG-rated firms generally have a higher market capitalization relative to low ESG-rated firms.

The HML factor is significant for the value-weighted high portfolio and all high-low portfolio weightings but only positive for the high- and high-low portfolios. Accordingly, the positive value premium indicates an increased orientation towards value firms among the high ESG-rated firms in our portfolio.

[34]Logically, this is also reflected in the intercept for the equal-weighted difference portfolio, although statistically insignificant.

[35]Except the equal-weighted high-low portfolio, which is significant at the 10% level.

The WML factor (i.e., the momentum premium) is positive for all coefficients except the equal-weighted high-low portfolio but only significant[36] for the value-weighted high portfolio. Accordingly, the positive momentum premium for the value-weighted high portfolio indicates that high ESG-rated firms are more tilted toward "winner" firms (i.e., the persistence in firm performance) than low ESG-rated firms.

Table 6.5: Carhart Four-Factor Model Quartile Regression Results

This table presents the Carhart Four-Factor model regression results for the Quartile portfolio. The intercept coefficients denote the abnormal monthly return (in percentage) for the equal- and value-weighted high-, low-, and high-low portfolios. The monthly risk-free rate (one-month U.S. treasury bill) is deducted from both the high- and low portfolios, while the *difference* portfolio (high-low) is constructed by subtracting the low portfolio returns from the returns of the high portfolio. The dependent variable is the monthly return ($r_{p,t}$) for each respective portfolio, and the independent variables represent the risk factors MktRF, SMB, HML, and WML. Heteroskedasticity Robust (HC1) Standard Errors are represented in parenthesis below each respective coefficient. The sample includes 48 monthly time series observations in the period January 2018 to 2021.

	Dependent Variable: Portfolio Return					
	High		**Low**		**High-Low**	
	EW	VW	EW	VW	EW	VW
Intercept	1.006***	0.287	1.045**	0.607***	-0.039	-0.320
	(0.284)	(0.188)	(0.463)	(0.202)	(0.317)	(0.250)
MktRF	1.041***	0.570***	0.973***	0.406***	0.069	0.164***
	(0.035)	(0.045)	(0.069)	(0.033)	(0.077)	(0.050)
SMB	0.358**	-0.371***	0.269	0.108	0.089	-0.479***
	(0.145)	(0.123)	(0.186)	(0.089)	(0.161)	(0.147)
HML	0.345**	0.126**	-0.206	-0.108	0.551***	0.233***
	(0.137)	(0.064)	(0.201)	(0.081)	(0.140)	(0.090)
WML	-0.051	0.126*	0.241	0.067	-0.292	0.059
	(0.132)	(0.070)	(0.231)	(0.092)	(0.180)	(0.111)
Observations	48	48	48	48	48	48
R^2	0.947	0.888	0.850	0.800	0.596	0.409
Adjusted R^2	0.943	0.878	0.836	0.781	0.558	0.354
Note:					*p<0.1; **p<0.05; ***p<0.01	

Table 6.5 presents the Carhart Four-Factor model regression results when considering Quartile. Overall, the obtained results using Quartile are largely similar to Decile in terms of statistical significance, although we observe changes to the magnitude of the coefficients.

When examining Quartile, the HML factor remains positive and significant for all portfolios

[36]Significant at the 5% level.

and weightings relative to Decile, except the equal-weighted high portfolio, which is significant at the 5% level. Neither WML coefficients are significant when considering Quartile, since the momentum premium is no longer significant at the 5% level for the value-weighted high portfolio.

6.2.3 Fama-French Five-Factor Model

Table 6.6: Fama-French Five-Factor Model Decile Regression Results

This table presents the Fama-French Five-Factor model regression results for the Decile portfolio. The intercept coefficients denote the abnormal monthly return (in percentage) for the equal- and value-weighted high-, low-, and high-low portfolios. The monthly risk-free rate (one-month U.S. treasury bill) is deducted from both the high- and low- portfolios, while the *difference* portfolio (high-low) is constructed by subtracting the low portfolio returns from the returns of the high portfolio. The dependent variable is the monthly return ($r_{p,t}$) for each respective portfolio, and the independent variables represent the risk factors MktRF, SMB, HML, RMW, and CMA. Heteroskedasticity Robust (HC1) Standard Errors are represented in parenthesis below each respective coefficient. The sample includes 48 monthly time series observations in the period January 2018 to 2021.

	Dependent Variable: Portfolio Return					
	High		**Low**		**High-Low**	
	EW	VW	EW	VW	EW	VW
Intercept	1.652***	0.714***	1.808***	1.030***	-0.156	-0.316
	(0.312)	(0.252)	(0.447)	(0.206)	(0.397)	(0.295)
MktRF	1.118***	0.751***	0.926***	0.332***	0.192**	0.419***
	(0.064)	(0.065)	(0.096)	(0.047)	(0.091)	(0.055)
SMB	0.727***	-0.264	0.734***	0.064	-0.007	-0.327*
	(0.222)	(0.182)	(0.272)	(0.119)	(0.252)	(0.194)
HML	0.159	-0.147	-0.540*	-0.247	0.699***	0.100
	(0.229)	(0.167)	(0.317)	(0.162)	(0.263)	(0.170)
RMW	0.433	0.387**	0.774**	0.139	-0.341	0.248
	(0.281)	(0.166)	(0.346)	(0.162)	(0.276)	(0.214)
CMA	0.345	0.671***	0.155	0.041	0.190	0.630***
	(0.279)	(0.235)	(0.383)	(0.209)	(0.324)	(0.202)
Observations	48	48	48	48	48	48
R^2	0.891	0.841	0.719	0.592	0.556	0.623
Adjusted R^2	0.878	0.822	0.686	0.544	0.503	0.578
Note:						*p<0.1; **p<0.05; ***p<0.01

When examining the Decile regression results for the Fama-French Five-Factor model presented in Table 6.6, we observe positive and significant alphas at the 1% level for all weightings in both high- and low portfolios when controlling for the RMW and CMA factors. As such, the positive intercepts of the equal-weighted high (low) portfolio indicate

that high (low) ESG-rated firms generate monthly abnormal returns of 1.652% (1.808%) in excess of the risk-free rate. Meanwhile, their value-weighted counterpart generates approximately half of the abnormal returns for the high- and low portfolios. However, we observe no significance at any level for either weighting of the high-low portfolio, which indicates that the difference in ESG performance is unable to explain the variation in returns using this model specification.

The MktRF factor is positive and significant at the 1% level for all portfolios and weightings[37]. Moreover, the market risk premium is greatest for the high portfolio regardless of weightings, suggesting that the high ESG-rated firms generate a greater market risk-adjusted return relative to low ESG-rated firms.

The SMB factor is only positive and statistically significant for the equal-weighted high- and low portfolios at the 1% level. This is similar to the Three-Factor model and indicates that both the high- and low portfolios have a greater proportion of small-capitalization firms.

For the Five-Factor model, the HML factor is only significant (although less pronounced) for the equal-weighted high-low portfolio compared to the Three- and Four-Factor models. Nevertheless, the positive coefficients indicate an increased orientation towards value firms among the high ESG-rated firms.

The RMW factor (i.e., the profitability premium) is positive and significant for the value-weighted (equal-weighted) high (low) portfolio at the 5% level. When value-weighted, the positive profitability premium indicates that high ESG-rated firms have more robust marked-adjusted profitability. Meanwhile, we identify no statistical significance at any weightings for the high-low difference portfolio.

The CMA factor (i.e., the investment premium) is positive and significant for the value-weighted high- and high-low portfolio at the 1% level. The positive investment premium coefficients suggest that high ESG-rated firms are more conservative in terms of capital investments.

[37]Except the equal-weighted high-low portfolio, which is significant at the 5% level.

Table 6.7: Fama-French Five-Factor Model Quartile Regression Results

This table presents the Fama-French Five-Factor model regression results for the Quartile portfolio. The intercept coefficients denote the abnormal monthly return (in percentage) for the equal- and value-weighted high-, low-, and high-low portfolios. The monthly risk-free rate (one-month U.S. treasury bill) is deducted from both the high- and low portfolios, while the *difference* portfolio (high-low) is constructed by subtracting the low portfolio returns from the returns of the high portfolio. The dependent variable is the monthly return ($r_{p,t}$) for each respective portfolio, and the independent variables represent the risk factors MktRF, SMB, HML, RMW, and CMA. Heteroskedasticity Robust (HC1) Standard Errors are represented in parenthesis below each respective coefficient. The sample includes 48 monthly time series observations in the period January 2018 to 2021.

	Dependent Variable: Portfolio Return					
	High		**Low**		**High-Low**	
	EW	VW	EW	VW	EW	VW
Intercept	0.952***	0.420**	1.251***	0.685***	-0.299	-0.265
	(0.189)	(0.171)	(0.308)	(0.153)	(0.278)	(0.218)
MktRF	1.094***	0.575***	0.966***	0.382***	0.128*	0.193***
	(0.036)	(0.039)	(0.066)	(0.034)	(0.072)	(0.045)
SMB	0.415***	-0.316**	0.375**	0.095	0.040	-0.411***
	(0.129)	(0.143)	(0.179)	(0.085)	(0.155)	(0.153)
HML	0.112	-0.012	-0.269	-0.059	0.382**	0.046
	(0.142)	(0.105)	(0.199)	(0.109)	(0.189)	(0.125)
RMW	0.157	0.148	0.489**	-0.067	-0.332*	0.215
	(0.173)	(0.116)	(0.197)	(0.106)	(0.192)	(0.158)
CMA	0.427***	0.122	-0.021	-0.150	0.448**	0.272
	(0.162)	(0.125)	(0.234)	(0.147)	(0.220)	(0.173)
Observations	48	48	48	48	48	48
R^2	0.954	0.887	0.855	0.803	0.616	0.445
Adjusted R^2	0.949	0.874	0.837	0.779	0.570	0.379
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 6.7 presents the Fama-French Five-Factor model regression results when considering Quartile. Overall, the obtained results using Quartile are moderately similar to Decile in terms of statistical significance, although we observe changes to the magnitude of the coefficients.

When examining Quartile, the MktRF factor remains positive and statistically significant for all portfolios and weightings, except for the equal-weighted high-low portfolio. Meanwhile, for the SMB factor, both the equal-weighted high- and low portfolios remain significant, while we now also observe significance for the equal-weighted high- and high-low portfolios. Moreover, as with Decile, the HML factor remains positive and significant (although weaker) for only the high-low portfolio. Similarly, the RMW factor only remains

significant, relative to Decile, for the equal-weighted low portfolio. Oppositely, the CMA factor is no longer significant for the value-weighted high- and high-low portfolios but for its value-weighted counterparts when considering Quartile.

6.2.4 Fama-French Five-Factor Model with Momentum

Table 6.8: Fama-French Five-Factor Model with Momentum Decile Regression Results

This table presents the Fama-French Five-Factor with Momentum model regression results for the Decile portfolio. The intercept coefficients denote the abnormal monthly return (in percentage) for the equal- and value-weighted high-, low-, and high-low portfolios. The monthly risk-free rate (one-month U.S. treasury bill) is deducted from both the high- and low portfolios, while the *difference* portfolio (high-low) is constructed by subtracting the low portfolio returns from the returns of the high portfolio. The dependent variable is the monthly return ($r_{p,t}$) for each respective portfolio, and the independent variables represent the risk factors MktRF, SMB, HML, RMW, CMA, and WML. Heteroskedasticity Robust (HC1) Standard Errors are represented in parenthesis below each respective coefficient. The sample includes 48 monthly time series observations in the period January 2018 to 2021.

	Dependent Variable: Portfolio Return					
	High		**Low**		**High-Low**	
	EW	VW	EW	VW	EW	VW
Intercept	1.357***	0.299	1.366**	0.790***	-0.010	-0.491
	(0.438)	(0.261)	(0.610)	(0.273)	(0.452)	(0.335)
MktRF	1.153***	0.800***	0.979***	0.360***	0.175*	0.440***
	(0.062)	(0.066)	(0.096)	(0.049)	(0.098)	(0.061)
SMB	0.686***	-0.320*	0.674**	0.031	0.013	-0.351*
	(0.217)	(0.169)	(0.272)	(0.120)	(0.255)	(0.199)
HML	0.245	-0.026	-0.411	-0.177	0.656**	0.151
	(0.244)	(0.143)	(0.338)	(0.159)	(0.271)	(0.179)
RMW	0.403	0.344**	0.729**	0.114	-0.326	0.230
	(0.263)	(0.158)	(0.323)	(0.152)	(0.277)	(0.224)
CMA	0.453*	0.822***	0.316	0.128	0.137	0.694***
	(0.263)	(0.214)	(0.370)	(0.203)	(0.336)	(0.212)
WML	0.265	0.372***	0.397	0.215	-0.132	0.157
	(0.217)	(0.114)	(0.307)	(0.139)	(0.214)	(0.134)
Observations	48	48	48	48	48	48
R^2	0.897	0.868	0.737	0.624	0.560	0.631
Adjusted R^2	0.882	0.848	0.698	0.570	0.496	0.577
Note:					*p<0.1; **p<0.05; ***p<0.01	

When examining the Decile regression results for the Fama-French Five-Factor model with Momentum (i.e., the WML factor) presented in Table 6.8, we observe positive and significant intercepts for all weightings for the low portfolio but only the equal-weighted

high portfolio. Hence, when equal-weighted (and value-weighted), the positive intercepts of the high (low) portfolio indicate that high (low) ESG-rated firms generate monthly abnormal returns of 1.357% (0.790%) in excess of the risk-free rate. Both coefficients' magnitude is less than the Fama-French Five-Factor model Decile results. In line with the findings from the Four-Factor model specification, which also controls for momentum, the difference in monthly abnormal returns between the equal-weighted high- and low portfolios is close to zero. However, we observe no significance at any level for either the equal- or value-weighted high-low portfolio, thus indicating that the difference in ESG performance is unable to explain the variation in returns using this model specification.

Regarding the market risk premium, we observe positive and significant coefficients at the 1% level for all portfolios and weightings, except the equal-weighted high-low portfolio. Furthermore, consistent with the results of the Decile Three- and Five-Factor models, the market risk premium is greatest for the high portfolio regardless of weightings when controlling for the momentum factor. This implies that the high ESG-rated firms generated a greater market risk-adjusted return than the low ESG-rated firms. Interestingly, adding the momentum factor results in higher magnitudes for all significant coefficients relative to all other Decile model specifications.

For the SMB factor, controlling for momentum does not notably impact the size premium and thus adds little explanation relative to the Decile Five-Factor model. Similarly, the HML factor is positive and remains significant (although weaker) for only the equal-weighted high-low portfolio at the 5% level. Moreover, the RMW factor remains positive and significant only for the value-weighted high portfolio and equal-weighted low portfolios at the 5% level. Finally, in line with the findings from the Five-Factor model, we still observe a positive and significant investment premium for the value-weighted high- and high-low portfolios. Accordingly, due to the similar results obtained from the Five-Factor model for the SMB, HML, RMW, and CMA factors, we do not elaborate any further on the effect of the size, value, and profitability premium.

Consistent with the findings using the Decile Four-Factor model specification, the WML factor is only positive and significant[38] for the value-weighted high portfolio. This indicates that our high portfolio is more tilted toward "winner" firms, i.e., that high ESG-rated

[38]However, now statistically significant at the 1% level.

firms show greater persistence in firm performance relative to the overall market.

Table 6.9: Fama-French Five-Factor Model with Momentum Quartile Regression Results

This table presents the Fama-French Five-Factor with Momentum model regression results for the Quartile portfolio. The intercept coefficients denote the abnormal monthly return (in percentage) for the equal- and value-weighted high-, low-, and high-low portfolios. The monthly risk-free rate (one-month U.S. treasury bill) is deducted from both the high- and low portfolios, while the *difference* portfolio (high-low) is constructed by subtracting the low portfolio returns from the returns of the high portfolio. The dependent variable is the monthly return ($r_{p,t}$) for each respective portfolio, and the independent variables represent the risk factors MktRF, SMB, HML, RMW, CMA, and WML. Heteroskedasticity Robust (HC1) Standard Errors are represented in parenthesis below each respective coefficient. The sample includes 48 monthly time series observations in the period January 2018 to 2021.

	Dependent Variable: Portfolio Return					
	High		**Low**		**High-Low**	
	EW	VW	EW	VW	EW	VW
Intercept	0.947***	0.258	1.003**	0.626***	-0.056	-0.369
	(0.283)	(0.185)	(0.459)	(0.210)	(0.332)	(0.249)
MktRF	1.094***	0.594***	0.995***	0.389***	0.099	0.205***
	(0.040)	(0.045)	(0.070)	(0.036)	(0.080)	(0.053)
SMB	0.415***	-0.338**	0.341*	0.087	0.073	-0.425***
	(0.139)	(0.136)	(0.188)	(0.090)	(0.158)	(0.155)
HML	0.114	0.035	-0.197	-0.042	0.311	0.076
	(0.159)	(0.100)	(0.230)	(0.118)	(0.206)	(0.129)
RMW	0.157	0.131	0.463**	-0.074	-0.307	0.204
	(0.175)	(0.115)	(0.188)	(0.105)	(0.189)	(0.160)
CMA	0.429**	0.181	0.069	-0.128	0.360	0.309*
	(0.173)	(0.115)	(0.222)	(0.149)	(0.229)	(0.165)
WML	0.005	0.146**	0.223	0.052	-0.218	0.093
	(0.133)	(0.073)	(0.226)	(0.099)	(0.181)	(0.123)
Observations	48	48	48	48	48	48
R^2	0.954	0.894	0.861	0.804	0.636	0.452
Adjusted R^2	0.948	0.879	0.840	0.776	0.583	0.372
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 6.9 shows the Fama-French Five-Factor model regression results when considering Quartile. Overall, the obtained results using Quartile are largely similar to Decile in terms of statistical significance, although we observe changes to the magnitude of the coefficients.

When examining Quartile and applying the Five-Factor model with momentum, we find the monthly abnormal returns to remain significant and positive, in line with the results for its Decile counterpart. Furthermore, the MktRF factor remains positive and

statistically significant for all portfolios and weightings, except for the equal-weighted high-low portfolio. Meanwhile, when examining Quartile, the SMB factor for the equal-weighted high portfolio remains positive and significant. However, we also observe negative and significant SMB coefficients for both the value-weighted high- and high-low portfolios.

On the contrary, we no longer observe any significant HML factor coefficients for any portfolios, regardless of weighting. Similarly, the RMW factor is no longer significant for the value-weighted high portfolio but remains significant for the equal-weighted low portfolio relative to its Decile counterpart. Finally, the CMA factor is now only significant for the equal-weighted high portfolio, while the WML factor remains significant (although weaker) for the value-weighted high portfolio.

7 Discussion

In this section, we discuss the most influential results described in Section 6 in relation to existing literature. In Section 2, we describe that there is a lack of research on the relationship between ESG performance and stock returns in emerging markets and that the existing studies on this topic have provided both inconsistent and inconclusive results (Atz et al., 2022; Friede et al., 2015; Griffin & Mahon, 1997). Accordingly, we aim to add to the current body of research by addressing our overarching research question, whether ESG performance significantly impacts the expected stock returns of firms in emerging markets[39] through discussing our findings in relation to the stated hypothesis, i.e., that investors pay a premium for holding high ESG-rated firms in emerging markets.

In addition, we attempt to expand our understanding of the supplementary research question by examining whether the restrictiveness of defining what constitutes high- and low ESG performance influences stock returns. Consequently, the following discussion attempts to broaden the current understanding of the influence of ESG performance on stock returns.

7.1 Portfolio Characteristics

From Figure 6.1, showing the cumulative portfolio returns, we observe that the Decile cutoff is, on average, more influenced by market fluctuations. A possible explanation could be that Decile is relatively less diversified than its Quartile counterpart, i.e., in terms of the number of firms included in the portfolio, as seen in the industrial sector composition. This is further substantiated by the descriptive portfolio results, shown in Table 6.1, which could be explained by the Quartile cutoff, on average, has higher market capitalization for each period compared to its Decile counterpart.

As presented in Section 6.1.3, the observation of the *Industrials* sector among the top three most occurring sectors for both cutoffs was in line with our initial expectations for emerging markets. The intuition is that emerging markets historically have been more labor-intensive (Hanson, 2012) due to a relatively large population combined with businesses that tend to avoid costly capital investments. Thus, we presumed emerging

[39] As presented in Section 3.

economies to be well represented in sectors where products can be cheaply produced (due to comparative advantages). Conversely, we were surprised to observe the *Financials* and *Information Technology* sectors among the top three most occurring sectors, as these are typically more capital-intensive. However, this could be explained by the recent technological advances and labor productivity observed in emerging economies, which could have mitigated some of the labor dependency.

7.2 Regression Results

The following part discusses the main findings from our regression results, emphasizing abnormal returns, exposure to the market premium, and exposure to the value and investment premium.

7.2.1 Abnormal Returns

As presented in Section 6.2, the results of our regressions show that although we consistently observe economically significant negative intercepts[40], none of our model specifications were able to identify statistically significant abnormal returns for the difference (high-low) portfolio. Moreover, the lack of statistically significant abnormal returns remains regardless of weightings and cutoffs, and the magnitude of the intercepts ranges between the interval of [-0.2%, -0.5%]. Accordingly, we do not find evidence that firm ESG performance (i.e., ESG rating) explains the difference in abnormal returns between firms with perceived good- and bad ESG practices (all else equal). Thus, we fail to reject our hypothesis based on the generated difference portfolio. Based on these results, we cannot infer the validity of our hypothesis, although we must retain it as a possibility simply because of the lack of evidence to reject it. On the contrary, we found both statistically and economically significant positive abnormal returns for each respective high- and low portfolio, irrespective of weightings and cutoffs.

Regarding our overarching research question, our findings suggest that both high- and low ESG-rated firms generate greater monthly abnormal returns[41], relative to the market,

[40]The value-weighted difference portfolio is economically significant across all model specifications, whereas its equal-weighted counterpart is economically insignificant (i.e., close to zero) for all specifications that control for momentum.

[41]In excess of the risk-free rate.

although the low ESG-rated firms also consistently outperform the high ESG-rated firms throughout our sample period. However, when adjusting for momentum, we observe that the monthly abnormal returns between the two portfolios even out, as reflected in the intercept of the equal-weighted difference portfolio. Yet, we do not see the same effect for its value-weighted counterpart. Meanwhile, for all models which include the momentum factor (WML), we observe consistently lower intercepts relative to their counterparts without momentum, which could suggest a negative relationship between monthly abnormal returns and momentum.

However, although divided, some of the existing research has established that higher ESG scores are frequently associated with higher expected returns (Kempf & Osthoff, 2007) and that the effect is particularly pronounced in emerging markets (Friede et al., 2015). In contrast to the findings of Kempf and Osthoff (2007), which established that high ESG-rated firms generate greater abnormal returns, we cannot detect a similarly significant relationship for our *difference* portfolio. Nevertheless, our findings showed significant abnormal returns for each respective high- and low portfolio. Hence, similar to the results of Atz et al. (2022), Renneboog et al. (2008), and Revelli and Viviani (2015), our findings fall under the inconclusive category[42]. However, it is crucial to be mindful of the distinction between the analyzed markets, whereas we limit our analysis to emerging markets, while the aforementioned research considers developed markets. As touched upon in Section 2.7, there are distinct differences between developed and emerging markets, which could explain some of the differences in our results.

When examining the effect of restrictiveness on the portfolio, we compared the results obtained using the Decile cutoff to its Quartile counterpart. For Decile, we consistently identified greater magnitudes of positive abnormal returns for the high- and low portfolios, regardless of weightings. Meanwhile, for the Decile high-low portfolio, we observe the opposite relationship where equal-weighted is consistently lower than its Quartile counterpart and vice-versa when value-weighted. This implies that when being more restrictive in what defines a good- or bad ESG performer, we obtain greater abnormal returns relative to the market. Accordingly, we find some evidence supporting our supplementary research question, where the relative difference in ESG scores between

[42]As highlighted in the discussion of positive, negative and inconclusive evidence of ESG scores and financial performance in Section 2.6.

high- and low ESG-rated firms influences the size of the premium paid by investors for holding such firms.

Overall, our findings imply greater abnormal market-adjusted returns when holding a portfolio composed of only low ESG-rated firms relative to an equally restrictive portfolio of only high ESG-rated firms. Nonetheless, this observation seems somewhat counterintuitive, given the current investment climate and growing focus on ESG considerations (Global Sustainable Investment Alliance, 2021). Moreover, as described in Section 2.1.2, when stakeholder preferences align with those of the firm, the intuition is that this affects the firm's financial performance positively by enhancing firm reputation and loyalty (H. B. Christensen et al., 2019, p. 48). If we assume that stakeholders are generally ESG-conscious, we observe the opposite of what this theory predicts. However, this is likely to be a wrongful assumption in light of our findings. This mechanism could then possibly be derived from an increased risk premium demanded by investors for holding a portfolio composed of firms with higher associated risk due to, e.g., moral, ethical, or environmental controversies.

7.2.2 Exposure to the Market Premium

When considering the market premium, we observe a consistently positive and significant relationship for all portfolios, irrespective of weightings and model specifications, except for the equal-weighted high-low portfolio[43]. The positive (and relatively higher) market premium implies that high ESG-rated firms have higher associated market-adjusted risk than low ESG-rated firms. This contradicts the intuition stated in our hypothesis that high ESG-rated firms are perceived to be less risky than low ESG-rated firms.

However, a potentially influential factor is that our ESG scores do not account for ESG controversies, as touched upon in Section 4.1.2. Consequently, there might be unobserved variables that may capture the unexplained associated risk for the high ESG-rated firms in our portfolios. On the other hand, another possible explanation is the differences in the GICS sector composition of the high- and low portfolios, as presented in Section 6.1.3. From the portfolio sector composition plots, shown in Figure 6.2, we observe that the composition of the high portfolio is tilted more toward industries that are typically more

[43]In particular, we only observe a significant market premium for the Three- and Five-Factor Decile model specifications.

sensitive to monetary policy changes and market turmoil (e.g., Information Technology and Financials). Our selected sample period covers approximately two years of the COVID-19 pandemic, where the sudden disruption to global demand and supply chains resulted in, e.g., a drastic global expansionary monetary policy with unparalleled quantitative easings to alleviate the burdens on the real economy[44]. Examining to which extent the COVID-19 pandemic impacted emerging markets is, however, outside the scope of this thesis and will not be elaborated on any further.

The explanatory variables in our models might not entirely suffice in explaining all relevant effects which influence the dependent variable. Thus, our results might be affected by omitted variable bias (Wooldridge, 2020), which implies that there are other factors that explain the dependent variable, i.e., portfolio return. Additionally, if our error term in the OLS regressions contains variables that are correlated with one of the explanatory factors, we have omitted variables that ultimately lead to biased estimators.

7.2.3 Exposure to the Size Premium

Our findings indicate a consistently negative and significant exposure to small-cap firms for the value-weighted high-low portfolio, across all models and irrespective of cutoffs, except for the Decile Five-Factor model specification with and without Momentum. Overall, this suggests that for our value-weighted difference portfolio, high ESG-rated firms are generally larger in firm market capitalization than low ESG-rated firms.

This is a logical result due to the size of the firm (i.e., the market capitalization) and the business area in which they operate; larger firms are generally subject to more rigid regulations and are consequently compelled to a greater degree of organizational transparency. Additionally, larger firms are under more frequent and organic scrutiny by investors, as evident by a higher trading volume (relative to smaller firms). By extension, consistent with the arguments presented by Sharma (2000) and Duque-Grisales and Aguilera-Caracuel (2021), it is reasonable to assume that larger firms also have more available capital and resources to incorporate ESG considerations into their value chain.

[44]The extreme inflows of funds to the world's leading economies, due to the various monetary stimulus packages designed to stimulate supply and demand resulted in, e.g., artificially inflated asset prices.

7.2.4 Exposure to the Value- and Investment Premium

In terms of the value premium, the difference portfolio predominantly shows a consistently significant and positive relationship irrespective of weighting and model specification[45]. Meanwhile, the investment premium is positive and significant for the Five-Factor model specification with and without Momentum for the Decile cutoff and the equal-weighted Quartile Five-Factor without Momentum.

As such, the magnitude of the value premium implies that there are more value firms than growth firms among the high ESG-rated firms in our portfolios. Moreover, the magnitude of the investment premium suggests that the high ESG-rated firms are more conservative (in terms of investment strategy) than the low ESG-rated firms. Thus, it is reasonable to assume that conservative value firms typically are larger firms with sufficient resources for incorporating ESG consideration. Hence, these findings are consistent with the discussion of the size premium in Section 7.2.3. Furthermore, consistent with the arguments of Sharma (2000), a feasible explanation could be that firms in emerging markets are more exposed to a lack of financial flexibility due to the scarcity of resources. Hence, executives are likely to pursue more profitable operational activities rather than prioritizing ESG initiatives deemed too expensive.

[45]Except the equal-weighted Five-Factor with Momentum irrespective of cutoffs, and both the value-weighted (and equal-weighted) Five-Factor with and without Momentum, for the Decile (and Quartile) cutoffs.

8 Conclusion

The purpose of our thesis was to study the effect of ESG performance on financial performance in emerging markets. To investigate this, we hypothesized that investors pay a premium for holding high ESG-rated firms over low ESG-rated firms in emerging markets. To test our hypothesis, we constructed two portfolios of high- and low ESG-rated firms, respectively, and applied a long-short strategy to generate a high-low portfolio. Then, we proceeded to examine historical stock returns in the four years between 2018-2021. We observed economically and statistically significant results for both the high- and low portfolio, where the low ESG-rated portfolio had higher monthly abnormal returns than the high ESG-rated portfolio (in excess of the market return). However, we did not observe any significant monthly abnormal returns for the high-low difference portfolio. Thus, we cannot conclusively reject our hypothesis of whether the observed premium captured by the difference portfolio, in fact, could be explained by differences in firm ESG performance between the high- and low portfolio.

Consequently, our models cannot infer causal relationships between the dependent and explanatory variables. Moreover, our results also point to high ESG-rated firms being more volatile than low ESG-rated firms. Thus, the paradox is how high ESG-rated firms are associated with both lower expected returns and higher expected volatility. We argue that unobserved explanatory variables might influence the dependent variable that is not captured by any of our factors.

As such, our results are added to the existing inconclusive evidence on the relationship between ESG performance and financial performance. However, our thesis still contributes toward a better understanding of the underresearched area of the influence of ESG on financial performance in emerging markets. Moreover, we captured the high- and low ESG-rated firms' performance during a period of high market volatility. Hence, this provided us with valuable insight into the potential driving forces behind their performance.

Regarding the supplementary research question, we found that the Decile cutoff is, on average, less diversified with increased volatility and yielded more extreme magnitudes for our factor loadings, i.e., economically higher monthly abnormal returns. Thus, this substantiates our supplementary research question and suggests that applying a more

restrictive criterion for good- or bad ESG performance leads to higher monthly abnormal returns compared to the market. Moreover, the high ESG-rated portfolio, on average, had a higher market value than the low ESG-rated portfolio, particularly evident for the Decile cutoff. This indicates that the top Decile of high ESG-rated firms tends to have a higher market capitalization in our respective sample period.

8.1 Limitations of our Thesis

Generally, conducting empirical analysis necessitates some underlying assumptions. To this extent, we have been conscious of our empirical choices while also stating our assumptions throughout this thesis. Despite our efforts to avoid introducing potential biases, our analysis is subject to inherent limitations. In terms of the sample size, one could argue that we would likely obtain more robust results by using a larger sample size to better reflect the overall development in emerging markets and increase the sampled time period. However, this proved somewhat difficult to accomplish due to the varying availability of ESG scores from our accessible providers. The same restrictions apply across our sample in the cross-section and time series.

8.1.1 Transaction Costs

For simplicity, we have ignored all potential consequences associated with transaction costs in our analysis. As portfolio rebalancing incurs transaction costs, it subsequently reduces the returns more than our data reflects. For instance, for our high-low portfolio, it is reasonable to assume that our constructed portfolio incurs miscellaneous transaction costs when holding both long- and short positions. Thus, this must be considered a limitation of our thesis.

8.1.2 Survivorship Bias

To address potential survivorship bias issues, we did not require firms to have ESG scores for the entire period. To this extent, we utilize all available data for the given period, irrespective of the firm's prior ESG performance. Moreover, the decision of whether firms are included in our constructed high- and low portfolios are objectively determined by the respective restrictiveness. Although we are not particularly concerned with survivorship-

biased data, we cannot completely eliminate the possibility of its existence.

8.2 Suggestions for Further Research

Our analysis provides valuable insight into an underresearched field in emerging markets, as most research on the link between ESG and financial performance is predominantly performed in developed markets. Our findings clearly present a link between ESG- and financial performance, where holding high ESG-rated firms leads to relatively lower expected returns. However, due to our results offering conflicting evidence, it remains challenging to provide a clear conclusion on the underlying causes of the observed effect. To this extent, further research is needed to identify the unobserved effects we have described. Moreover, it would be interesting to further examine whether ESG should be included in future asset pricing models.

Lastly, we have exclusively downloaded ESG information regarding stocks listed in emerging markets. Therefore, it would have been interesting to explore whether our results hold for other securities in emerging markets with ESG information available, e.g., fixed-income securities. However, ultimately, we deemed this beyond the scope of our research questions and therefore refrained from going into further details in our thesis.

References

Ambec, S., & Lanoie, P. (2008). Does it pay to be green? a systematic overview. *The Academy of Management Perspectives*, 45–62.

Amel-Zadeh, A., & Serafeim, G. (2018). Why and how investors use esg information: Evidence from a global survey. *Financial Analysts Journal*, *74*(3), 87–103. https://doi.org/10.2469/faj.v74.n3.2

Atz, U., Van Holt, T., Liu, Z. Z., & Bruno, C. C. (2022). Does sustainability generate better financial performance? review, meta-analysis, and propositions. *Journal of Sustainable Finance & Investment*, 1–24. https://doi.org/10.1080/20430795.2022.2106934

Auer, B. R. (2016). Do socially responsible investment policies add or destroy european stock portfolio value? *Journal of Business Ethics*, *135*(2), 381–397. https://doi.org/10.1007/s10551-014-2454-7

Basar, S. (2021). Transparency, methodology, and consistency in esg scoring. Retrieved November 8, 2022, from https://www.bloomberg.com/professional/blog/transparency-methodology-and-consistency-in-esg-scoring/

Berg, F., Koelbel, J. F., & Rigobon, R. (2019). Aggregate confusion: The divergence of esg ratings. *Forthcoming Review of Finance*. https://doi.org/10.1093/rof/rfac033

Berry, T. C., & Junkus, J. C. (2013). Socially responsible investing: An investor perspective. *Journal of Business Ethics*, *112*(4), 707–720. https://doi.org/10.1007/s10551-012-1567-0

Bloomberg. (2022a). Environmental, social & governance (esg) bloomberg professional solutions. Retrieved November 4, 2022, from https://www.bloomberg.com/professional/dataset/global-environmental-social-governance-data/

Bloomberg. (2022b). *Esg analysis help page* (Report).

Camilleri, M. A. (2020). The market for socially responsible investing: A review of the developments. *Social Responsibility Journal*, *17*(3), 412–428. https://doi.org/10.1108/SRJ-06-2019-0194

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, *52*(1), 57–82. https://doi.org/10.1111/j.1540-6261.1997.tb03808.x

Chakravarthy, J., DeHaan, E., & Rajgopal, S. (2014). Reputation repair after a serious restatement. *The Accounting Review*, *89*(4), 1329–1363. https://doi.org/10.2308/accr-50716

Chatterji, A. K., & Toffel, M. W. (2010). How firms respond to being rated. *Strategic Management Journal*, *31*(9), 917–945. https://doi.org/10.1002/smj.840

Christensen, D. M., Serafeim, G., & Sikochi, A. (2022). Why is corporate virtue in the eye of the beholder? the case of esg ratings. *The Accounting Review*, *97*(1), 147–175. https://doi.org/10.2308/TAR-2019-0506

Christensen, H. B., Hail, L., & Leuz, C. (2019). *Adoption of csr and sustainability reporting standards: Economic analysis and review* (Vol. 623). National Bureau of Economic Research Cambridge, MA, USA.

Clark, G. L., Feiner, A., & Viehs, M. (2015). From the stockholder to the stakeholder: How sustainability can drive financial outperformance. https://doi.org/10.2139/ssrn.2508281

Danielson, M. G., Heck, J. L., & Shaffer, D. (2008). Shareholder theory-how opponents and proponents both get it wrong. *Available at SSRN 1309066*. https://doi.org/10.2139/ssrn.1309066

Daugaard, D. (2020). Emerging new themes in environmental, social and governance investing: A systematic literature review. *Accounting & Finance*, *60*(2), 1501–1530. https://doi.org/10.1111/acfi.12479

Diamonte, R. L., Liew, J. M., & Stevens, R. L. (1996). Political risk in emerging and developed markets. *Financial Analysts Journal*, *52*(3), 71–76. https://doi.org/10.2469/faj.v52.n3.1998

Duque-Grisales, E., & Aguilera-Caracuel, J. (2021). Environmental, social and governance (esg) scores and financial performance of multilatinas: Moderating effects of geographic international diversification and financial slack. *Journal of Business Ethics*, *168*(2), 315–334. https://doi.org/10.1007/s10551-019-04177-w

Eccles, R. G., Ioannou, I., & Serafeim, G. (2014). The impact of corporate sustainability on organizational processes and performance. *Management Science*, *60*(11), 2835–2857. https://doi.org/10.1287/mnsc.2014.1984

Eccles, R. G., Serafeim, G., Seth, D., & Ming, C. C. Y. (2013). The performance frontier: Innovating for a sustainable strategy: Interaction. *Harvard Business Review*, *91*(7), 17–18.

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *The Journal of Finance*, *47*(2), 427–465. https://doi.org/10.1111/j.1540-6261.1992.tb04398.x

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

Freeman, R. E. (2001). A stakeholder theory of the modern corporation. *Perspectives in Business Ethics Sie*, *3*(144), 38–48.

Freeman, R. E., & Phillips, R. A. (2002). Stakeholder theory: A libertarian defense. *Business Ethics Quarterly*, *12*(3), 331–349.

Freeman, R. E., & Reed, D. L. (1983). Stockholders and stakeholders: A new perspective on corporate governance. *California Management Review*, *25*(3), 88–106.

French, K. R. (2022). Description of fama/french 5 factors for emerging markets. Retrieved September 3, 2022, from http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/f-f_5emerging.html

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

Friedman, M. (1970). A friedman doctrine: The social responsibility of business is to increase its profits. *The New York Times Magazine*, *13*(1970), 32–33.

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. https://doi.org/10.1016/j.jclepro.2017.02.180

Gibson Brandon, R., Krueger, P., & Schmidt, P. S. (2021). Esg rating disagreement and stock returns. *Financial Analysts Journal*, *77*(4), 104–127. https://doi.org/10.1080/0015198X.2021.1963186

Global Sustainable Investment Alliance. (2021). Global sustainable investment review 2020. http://www.gsi-alliance.org

Griffin, J. J., & Mahon, J. F. (1997). The corporate social performance and corporate financial performance debate: Twenty-five years of incomparable research. *Business & Society*, *36*(1), 5–31. https://doi.org/10.1177/000765039703600102

Grohmann, A. (2018). Financial literacy and financial behavior: Evidence from the emerging asian middle class. *Pacific-Basin Finance Journal*, *48*, 129–143. https://doi.org/10.1016/j.pacfin.2018.01.007

Hanson, G. H. (2012). The rise of middle kingdoms: Emerging economies in global trade. *Journal of Economic Perspectives*, *26*(2), 41–64. https://doi.org/10.1257/jep.26.2.41

Huang, X. B., & Watson, L. (2015). Corporate social responsibility research in accounting. *Journal of Accounting Literature*, *34*(1), 1–16. https://doi.org/10.1016/j.acclit.2015.03.001

Kempf, A., & Osthoff, P. (2007). The effect of socially responsible investing on portfolio performance. *European Financial Management*, *13*(5), 908–922. https://doi.org/10.1111/j.1468-036X.2007.00402.x

Kitzmueller, M., & Shimshack, J. (2012). Economic perspectives on corporate social responsibility. *Journal of Economic Literature*, *50*(1), 51–84. https://doi.org/10.1257/jel.50.1.51

McWilliams, A., & Siegel, D. (2001). Corporate social responsibility: A theory of the firm perspective. *Academy of Management Review*, *26*(1), 117–127. https://doi.org/10.5465/amr.2001.4011987

Mendelsohn, R., Dinar, A., & Williams, L. (2006). The distributional impact of climate change on rich and poor countries. *Environment and Development Economics*, *11*(2), 159–178. https://doi.org/10.1017/S1355770X05002755

Mendes-Da-Silva, W., Onusic, L. M., & Bergmann, D. R. (2014). The influence of e-disclosure on the ex-ante cost of capital of listed companies in brazil. *Journal of Emerging Market Finance*, *13*(3), 335–365. https://doi.org/10.1177/0972652714550928

Miralles-Quirós, M. M., Miralles-Quirós, J. L., & Valente Gonçalves, L. M. (2018). The value relevance of environmental, social, and governance performance: The brazilian case. *Sustainability*, *10*(3), 574. https://doi.org/10.3390/su10030574

Morningstar. (2022). Total return - definition. Retrieved November 19, 2022, from https://gladmainnew.morningstar.com/home/GetDocumentLibDefinitions.aspx?type=DATAID&definitionId=HP010&dateformat=M/d/yyyy

MSCI. (2020). Global industry classification standard (gics) methodology. Retrieved November 20, 2022, from https://www.msci.com/documents/1296102/11185224/GICS+Methodology+2020.pdf/9caadd09-790d-3d60-455b-2a1ed5d1e48c?t=1578405935658

MSCI. (2022). Msci emerging markets index (usd) factsheet. Retrieved November 4, 2022, from https://www.msci.com/documents/10199/c0db0a48-01f2-4ba9-ad01-226fd5678111

Nasdaq. (2019). Thomson reuters and blackstone in discussions with london stock exchange group regarding the refinitiv business. Retrieved November 9, 2022, from https://www.nasdaq.com/press-release/thomson-reuters-and-blackstone-in-discussions-with-london-stock-exchange-group

Nasdaq. (2020). Morningstar to acquire sustainalytics and expand access to esg research, data, and analytics for investors worldwide. Retrieved November 13, 2022, from https://www.nasdaq.com/press-release/morningstar-to-acquire-sustainalytics-and-expand-access-to-esg-research-data-and

Nasdaq. (2021). Thomson reuters announces closing of sale of refinitiv to london stock exchange group. Retrieved November 9, 2022, from https://www.nasdaq.com/press-

release/thomson-reuters-announces-closing-of-sale-of-refinitiv-to-london-stock-exchange-group

Pedersen, L. H., Fitzgibbons, S., & Pomorski, L. (2021). Responsible investing: The esg-efficient frontier. *Journal of Financial Economics*, *142*(2), 572–597. https://doi.org/10.1016/j.jfineco.2020.11.001

Porter, M. E., & van der Linde, C. (1995). Toward a new conception of the environment-competitiveness relationship. *Journal of Economic Perspectives*, *9*(4), 97–118. https://doi.org/10.1257/jep.9.4.97

Refinitiv. (2022a). About us - we are refinitiv. Retrieved November 9, 2022, from https://www.refinitiv.com/en/about-us

Refinitiv. (2022b). Sustainable finance. Retrieved November 10, 2022, from https://www.refinitiv.com/en/sustainable-finance

Refinitiv. (2022c). Environmental, social and governance scores from refinitiv. Retrieved November 9, 2022, from https://www.refinitiv.com/content/dam/marketing/en_us/documents/methodology/refinitiv-esg-scores-methodology.pdf

Renneboog, L., Ter Horst, J., & Zhang, C. (2008). Socially responsible investments: Institutional aspects, performance, and investor behavior. *Journal of Banking & Finance*, *32*(9), 1723–1742. https://doi.org/10.1016/j.jbankfin.2007.12.039

Revelli, C., & Viviani, J.-L. (2015). Financial performance of socially responsible investing (sri): What have we learned? a meta-analysis. *Business Ethics: A European Review*, *24*(2), 158–185. https://doi.org/10.1111/beer.12076

Ruf, B. M., Muralidhar, K., Brown, R. M., Janney, J. J., & Paul, K. (2001). An empirical investigation of the relationship between change in corporate social performance and financial performance: A stakeholder theory perspective. *Journal of Business Ethics*, *32*(2), 143–156. https://doi.org/10.1023/A:1010786912118

Sargis, M., & Wang, P. (2020). How does investing in esg companies affect returns? Retrieved October 21, 2022, from https://www.morningstar.com/insights/2020/02/19/esg-companies

Sharma, S. (2000). Managerial interpretations and organizational context as predictors of corporate choice of environmental strategy. *Academy of Management Journal*, *43*(4), 681–697. https://doi.org/10.5465/1556361

Smith, H. J. (2003). The shareholders vs. stakeholders debate. *MIT Sloan Management Review*, *44*(4), 85–90.

Starks, L. T., Venkat, P., & Zhu, Q. (2017). Corporate esg profiles and investor horizons. *University of Texas at Austin*. https://doi.org/10.2139/ssrn.3049943

Sustainalytics. (2021). Esg risk ratings - methodology abstract - version 2.1. Retrieved November 13, 2022, from https://connect.sustainalytics.com/hubfs/INV/Methodology/Sustainalytics_ESG%20Ratings_Methodology%20Abstract.pdf

Sustainalytics. (2022). Esg risk ratings. Retrieved November 13, 2022, from https://www.sustainalytics.com/corporate-solutions/esg-solutions/esg-risk-ratings

US SIF. (2020). Report on us sustainable and impact investing trends 2020. https://www.ussif.org/currentandpast

Vance, S. C. (1975). Are socially responsible corporations good investment risks. *Management Review*, *64*(8), 19–24.

van Duuren, E., Plantinga, A., & Scholtens, B. (2016). Esg integration and the investment management process: Fundamental investing reinvented. *Journal of Business Ethics*, *138*(3), 525–533. https://doi.org/10.1007/s10551-015-2610-8

Walley, N., & Whitehead, B. (1994). It's not easy being green. *Harvard Business Review*, *72*(3), 46–51. https://hbr.org/1994/05/its-not-easy-being-green

Wang, Z., Liao, K., & Zhang, Y. (2022). Does esg screening enhance or destroy stock portfolio value? evidence from china. *Emerging Markets Finance and Trade*, *58*(10), 2927–2941. https://doi.org/10.1080/1540496X.2021.2014317

Womack, K., & Zhang, Y. (2003). Understanding risk and return, the capm, and the fama-french three-factor model. *No. 03-111*. Retrieved November 11, 2022, from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=481881

Wooldridge, J. M. (2020). *Introductory econometrics: A modern approach* (7th ed.). Cengage.

Zehetmayr, M., & Brandau, N. (2021). How environmental, social and governance (esg) data providers compare. Retrieved October 14, 2022, from https://www.ey.com/en_bh/financial-services-emeia/how-environmental-social-and-governance-esg-data-providers-compare

Appendix

A1 Regression Model Testing

Table A1.1: Decile: Augmented Dickey-Fuller Test for Stationarity

These tables represent the results of the Decile (Table A1.1) and Quartile (Table A1.2) Augmented Dickey-Fuller test for stationarity for the equal- and value-weighted High-, Low-, and High-Low portfolios. The null hypothesis (H_0) for the Augmented Dickey-Fuller Test is non-stationarity. Moreover, the table describes the ADF statistic *(Statistic)*, p-value *(p-value)*, number of lags *(Lags)*, number of observations *N(obs)*, and the critical values of *1%*, *5%*, and *10%*, respectively. The null hypothesis is rejected for *p<5%*, which indicates stationarity.

	High		**Low**		**High-Low**	
	EW	VW	EW	VW	EW	VW
Statistic	-5.931	-6.384	-5.430	-5.065	-5.762	-6.910
p-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Lags	0.000	0.000	0.000	0.000	0.000	0.000
N(obs)	47.000	47.000	47.000	47.000	47.000	47.000
1%	-3.577	-3.577	-3.577	-3.577	-3.577	-3.577
5%	-2.925	-2.925	-2.925	-2.925	-2.925	-2.925
10%	-2.600	-2.600	-2.600	-2.600	-2.600	-2.600
p<5%	Stationary	Stationary	Stationary	Stationary	Stationary	Stationary

Table A1.2: Quartile: Augmented Dickey-Fuller Test for Stationarity

	High		**Low**		**High-Low**	
	EW	VW	EW	VW	EW	VW
Statistic	-6.089	-6.828	-5.825	-5.583	-5.400	-3.903
p-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)
Lags	0.000	0.000	0.000	0.000	0.000	7.000
N(obs)	47.000	47.000	47.000	47.000	47.000	40.000
1%	-3.577	-3.577	-3.577	-3.577	-3.577	-3.605
5%	-2.925	-2.925	-2.925	-2.925	-2.925	-2.937
10%	-2.600	-2.600	-2.600	-2.600	-2.600	-2.606
p<5%	Stationary	Stationary	Stationary	Stationary	Stationary	Stationary

Table A1.3: Decile: Breusch-Godfrey Test for Autocorrelation

These tables present the Decile (Table A1.3) and Quartile (Table A1.4) Fama-French results of the Breusch-Godfrey test for Autocorrelation, where *Chi-squared* values and associated *p-values* (in parenthesis) for the equal- and value-weighted High-, Low-, and High-Low portfolios. The null hypothesis (H_0) is that there is no autocorrelation. H_0 is rejected for p<5%, which indicates autocorrelation. The test is performed for all the following model specifications; Fama-French Three-Factor *(FF3)*; Carhart Four-Factor *(Carhart4)*; Fama-French Five-Factor *(FF5)*; Fama-French Five-Factor with Momentum *(FF5(Mom))*.

	High		**Low**		**High-Low**	
	EW	VW	EW	VW	EW	VW
FF3	1.339	0.694	1.030	1.124	1.415	0.463
	(0.253)	(0.710)	(0.436)	(0.372)	(0.219)	(0.889)
Carhart4	1.071	0.614	1.518	1.712	1.357	0.476
	(0.408)	(0.776)	(0.181)	(0.124)	(0.246)	(0.881)
FF5	1.065	1.156	1.490	1.300	2.210	0.477
	(0.413)	(0.354)	(0.192)	(0.274)	(0.047)	(0.880)
FF5(Mom)	0.877	1.515	1.995	1.932	1.994	0.507
	(0.555)	(0.185)	(0.073)	(0.083)	(0.073)	(0.859)

Table A1.4: Quartile: Breusch-Godfrey Test for Autocorrelation

	High		**Low**		**High-Low**	
	EW	VW	EW	VW	EW	VW
FF3	0.733	1.365	0.917	1.318	0.911	1.233
	(0.676)	(0.241)	(0.522)	(0.263)	(0.527)	(0.307)
Carhart4	0.748	2.018	0.980	1.199	0.818	1.216
	(0.663)	(0.068)	(0.474)	(0.328)	(0.604)	(0.318)
FF5	0.471	2.097	1.162	1.335	1.980	1.189
	(0.884)	(0.059)	(0.350)	(0.257)	(0.074)	(0.334)
FF5(Mom)	0.458	3.573	1.150	1.250	1.624	1.034
	(0.892)	(0.004)	(0.358)	(0.301)	(0.150)	(0.435)

Table A1.5: Decile: Breusch-Pagan Test for Heteroskedasticity

These tables present the Decile (Table A1.5) and Quartile (Table A1.6) results of the Breusch-Pagan test for Heteroskedasticity, with P-values in parentheses. The null hypothesis (H_0) is that we have homoskedasticity, i.e., that there is no heteroskedasticity. In the case of a P-value less than 5%, we reject the null hypothesis and conclude that heteroskedasticity is present in our model. The test is performed for all the following model specifications; Fama-French Three-Factor *(FF3)*; Carhart Four-Factor *(Carhart4)*; Fama-French Five-Factor *(FF5)*; Fama-French Five-Factor with Momentum *(FF5(Mom))*.

	High		**Low**		**High-Low**	
	EW	VW	EW	VW	EW	VW
FF3	3.614	3.771	0.660	0.625	6.148	0.368
	(0.306)	(0.287)	(0.882)	(0.891)	(0.105)	(0.947)
Carhart4	3.403	3.755	1.170	0.352	6.169	0.669
	(0.493)	(0.440)	(0.883)	(0.986)	(0.187)	(0.955)
FF5	4.223	4.641	2.344	3.079	5.692	0.855
	(0.518)	(0.461)	(0.800)	(0.688)	(0.337)	(0.973)
FF5(Mom)	3.900	4.591	2.141	1.764	5.719	1.076
	(0.690)	(0.597)	(0.906)	(0.940)	(0.455)	(0.983)

Table A1.6: Quartile: Breusch-Pagan Test for Heteroskedasticity

	High		**Low**		**High-Low**	
	EW	VW	EW	VW	EW	VW
FF3	2.030	2.634	0.744	1.492	4.605	0.778
	(0.566)	(0.452)	(0.863)	(0.684)	(0.203)	(0.855)
Carhart4	2.904	3.009	1.593	2.628	4.629	1.205
	(0.574)	(0.556)	(0.810)	(0.622)	(0.328)	(0.877)
FF5	4.000	3.959	2.638	1.421	5.603	2.072
	(0.549)	(0.555)	(0.756)	(0.922)	(0.347)	(0.839)
FF5(Mom)	5.470	4.212	3.506	3.060	3.604	2.598
	(0.485)	(0.648)	(0.743)	(0.801)	(0.730)	(0.857)

Table A1.7: Variance Inflation Factor test for Multicollinearity

This table presents the results of the Variance Inflation Factor test, where we observe that multicollinearity is not a problem for our data as VIF values above 10 indicate multicollinearity. Therefore, we do not violate OLS assumption ii). The test is performed using all the risk factors in the following model specifications; Fama-French Three-Factor *(FF3)*; Carhart Four-Factor *(Carhart4)*; Fama-French Five-Factor *(FF5)*; Fama-French Five-Factor with Momentum *(FF5(Mom))*.

	FF3	**Carhart4**	**FF5**	**FF5(Mom)**
Intercept	1.017	1.406	1.035	1.421
MktRF	1.041	1.114	1.282	1.427
SMB	1.056	1.081	1.121	1.140
HML	1.097	1.813	3.481	3.696
RMW			1.218	1.225
CMA			3.236	3.443
WML		1.842		1.978

A2 Regression Diagnostic Plots

We created diagnostic plots for all regressions conducted for our analysis. However, due to similar observed patterns when inspecting the diagnostic plots, we decided to include only a few representative plots. Thus, in the following, we present the interpretation of the diagnostic plots, as well as three different regression diagnostic plots for the Fama-French Five-Factor Model with momentum, as this model specification includes the most risk factors.

The *Residual vs. Fitted* plot is useful to determine if the residuals of the regression model exhibit non-linear patterns. We consider the residual to follow a linear pattern if they display a *roughly* horizontal red line.

The *Normal Q-Q* plot is useful to determine if the residuals of the regression model are normally distributed. We consider the residuals to be normally distributed if they *roughly* fall evenly along the straight diagonal line.

The *Scale-Location* plot is useful to determine the homoskedasticity among the residuals in our regression model. For example, if the displayed red line is *roughly* horizontal, we can assume that the residuals display equal variance.

The *Residuals vs. Leverage* plot is useful to determine outliers (i.e., influential observations) among the residuals in our regression model. If any observations fall *outside* of the red lines (which denotes the Cook's distance), we consider this to be an influential observation.

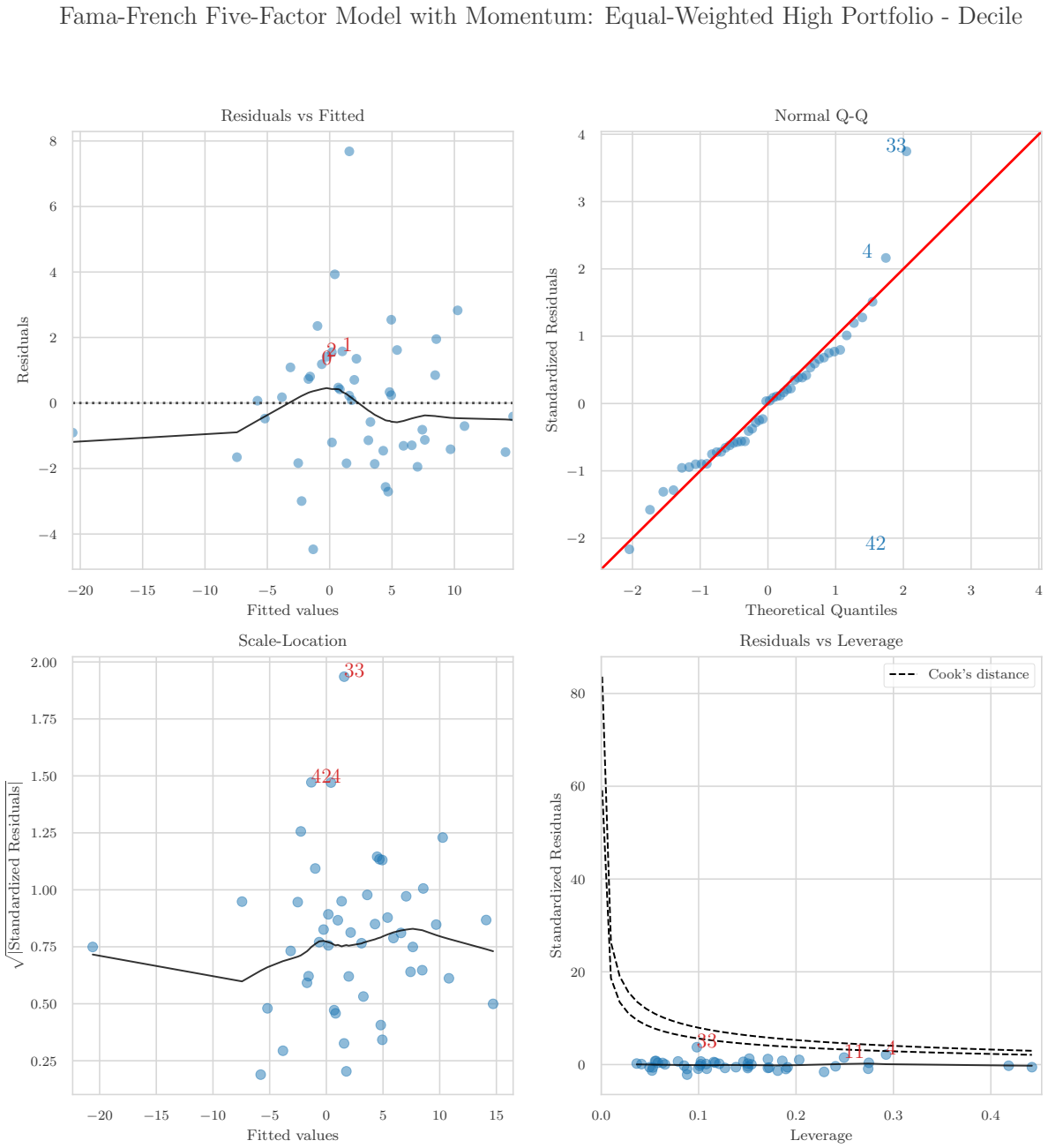

Figure A2.1: Equal-Weighted High Portfolio - Decile, when applying the Fama-French Five-Factor Model with Momentum

In the Equal-Weighted High Portfolio - Decile, when applying the Fama-French Five-Factor Model with Momentum, we observe that the *Residuals vs. Fitted* subplot lies on a roughly horizontal line. This suggests that a linear regression model is appropriate. Moreover, the observations land on a fairly straight diagonal line for the *Normal Q-Q* subplot and a roughly horizontal line for the *Scale-Location* subplot, despite the presence of some influential outliers. As such, from the *Residuals vs. Leverage* subplot, the observed outliers probably stem from the turbulent COVID-19 period. Overall, the diagnostic plots indicate that our residuals can confidently be deemed normally distributed and have equal variance.

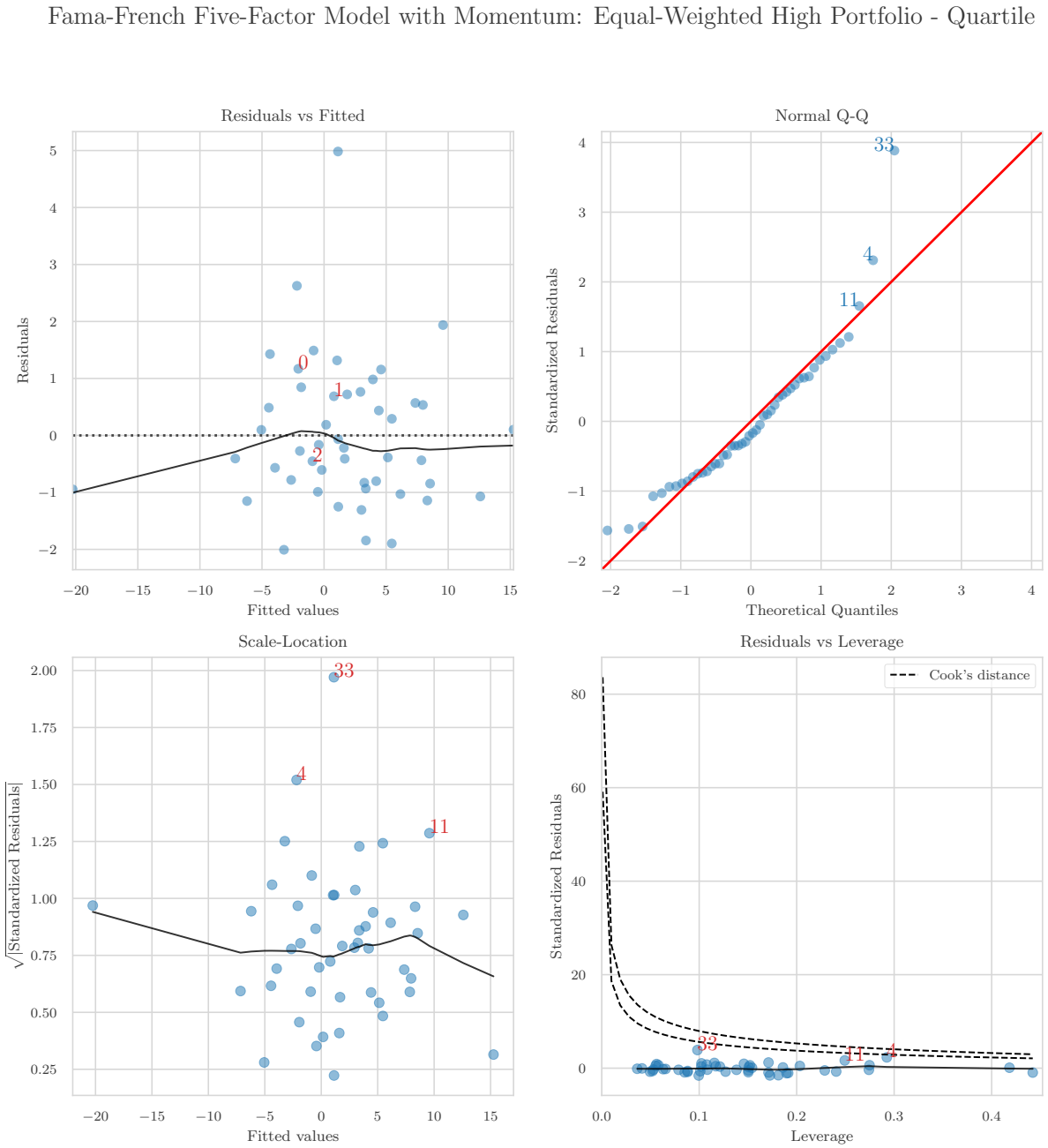

Figure A2.2: Equal-Weighted High Portfolio - Quartile, when applying the Fama-French Five-Factor Model with Momentum

For the *Residuals vs. Fitted* subplot, we observe a roughly horizontal line, which suggests that a linear regression model is appropriate for our Equal-Weighted High portfolio - Quartile. Moreover, the observations land on a fairly straight diagonal line for the *Normal Q-Q* subplot despite the presence of some influential outliers at each tail. Similarly, we observe that the observations fall on a roughly horizontal line in terms of the *Scale-Location* subplot. Overall, the diagnostic plots indicate that our residuals can confidently be deemed normally distributed and have equal variance. The same influential outliers, as presented in the Equal-Weighted High Portfolio - Decile, seem to be the likely explanation for our observed discrepancies.

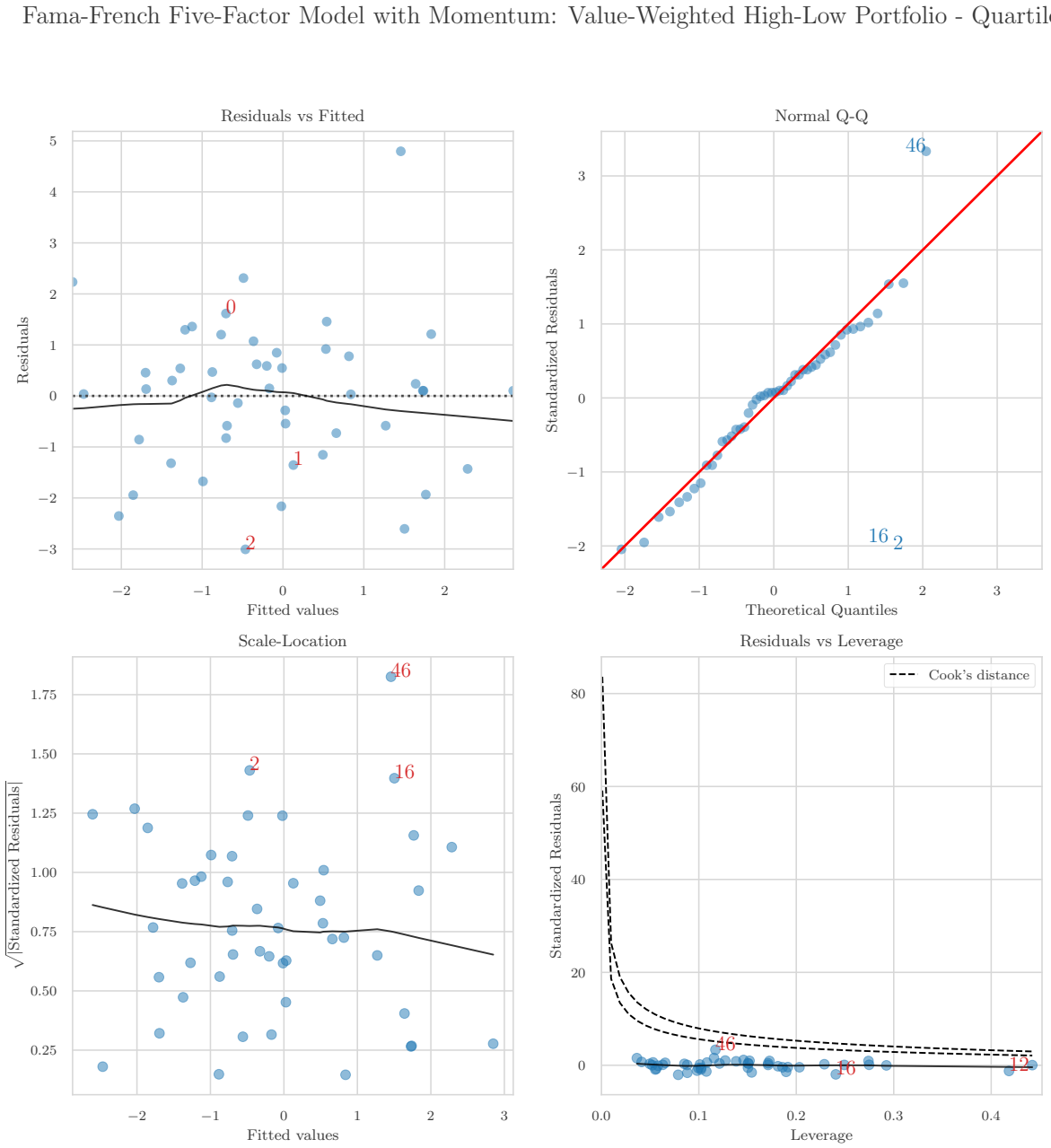

Figure A2.3: Value-Weighted High-Low Portfolio - Quartile, when applying the Fama-French Five-Factor Model with Momentum

For the *Residuals vs. Fitted* subplot, we observe a roughly horizontal line, which suggests that a linear regression model is appropriate for our Value-Weighted High-Low Portfolio - Quartile. Moreover, the observations land on a fairly straight diagonal line for the *Normal Q-Q* subplot and a roughly horizontal line for the *Scale-Location* subplot. This indicates that our residuals can be deemed normally distributed and have equal variance. It is noteworthy that the same influential outliers, as presented for both the Equal-Weighted High Portfolio - Decile and Equal-Weighted High Portfolio - Quartile, still seem to be relevant for the observed discrepancies.