

# MEMOIRE

Présenté en vue de l'obtention du Master en Ingénieur de gestion,  
finalité spécialisée

## ESG Integration in Equity Portfolios: An Empirical Study on Financial Performance across US and European Markets

Par Audric NEU AREND

Directeur: Mathias SCHMIT  
Assesseur: Bruno FARBER

Année académique 2023 - 2024

J'autorise la consultation de ce mémoire

## **Executive Summary**

Growing concerns about climate change and social disparities have driven a significant rise in the importance of sustainable finance and investing within the financial sector. This has led to substantial regulatory changes mandating companies to disclose sustainable information and meet specific criteria for classifying financial products. Consequently, the sustainable financial industry rapidly expands and provides diverse investment options.

Early adopters of sustainable investing were initially driven by personal values like ethics and environmental concerns. However, recent surveys indicate a shift in focus. Today, significant ESG holdings are held primarily by investors expecting market outperformance, viewing ESG (Environmental, Social, Governance) information as material for investment success. Nonetheless, some investors are cautious due to concerns about potential underperformance, highlighting the enduring significance of financial performance, even among those who prioritise making a positive impact through sustainable investing.

Most research emphasises the significance of ESG factors for companies, highlighting a positive relationship between sustainability and company financial performance. This relationship is often attributed to factors such as proactive environmental initiatives providing competitive advantages, strong social reputations improving stakeholder relations, and better alignment between shareholder and management interests. These findings align with established theories such as Stakeholder Theory, Porter Hypothesis, Natural Resource-Based View, and Agency Theory.

However, the research community lacks consensus on whether ESG investing strategies can generate abnormal returns for investors. Variability in findings is often attributed to factors such as data type, geographic scope, methodology, time periods, and other variables, leading to inconclusive results. According to the Efficient Market Hypothesis, there should be no link between risk-adjusted returns and ESG, as all publicly available and historical information, including third-party ESG ratings, should be incorporated into asset pricing, making consistent abnormal returns impossible. However, empirical evidence suggests that markets are not always perfectly efficient, and potential anomalies could arise from factors such as investor irrationality, biases, and tastes. In the context of ESG, such anomalies might emerge due to shifts in ESG awareness and motivation, or underreactions to ESG-related information. Other factors possibly influencing performance include reduced diversification opportunities from implementing ESG screens and higher turnover rates. However, some argue that these drawbacks are balanced out by the superior quality of the remaining assets. Overall, the empirical studies offer a mixed view, underlining the need for further investigation into integrating ESG factors into investment portfolios to understand to what extent it is possible to achieve financial success with sustainable portfolios.

This quantitative study examined to what extent the integration of ESG data into stock portfolio construction affects risk-adjusted returns and whether investors can devise strategies to leverage this information effectively. Portfolios were constructed from constituents of the S&P 500 and STOXX Europe 600 indices using a holding-based approach. The analysis utilised the Refinitiv ESG combined score (ESGC), which adjusts companies' ESG scores based on controversies, providing a holistic view of sustainability performance. To examine strategies beyond basic negative screening, which is often linked to underperformance due to restrictive criteria or timing costs, the portfolios included annually rebalanced traditional top and bottom ESGC portfolios, as well as innovative ESGC-weighted portfolios. The Fama-French 5-factor model was employed to assess risk-adjusted returns over 14 years (2010-2023), divided into two 7-year segments. A separate analysis examined the volatile period of the COVID-19 pandemic, encompassing both the drawdown and recovery phases.

Examining top and bottom ESGC score portfolios revealed statistically similar risk-adjusted returns over extended periods, regardless of whether investors pursued sustainable or unsustainable investment strategies. Variations in simple returns primarily stemmed from differing risk exposures. Sustainable portfolios typically exhibited lower systematic risk, favoured larger, value-oriented stocks, and adopted conservative investment approaches. However, statistically significant alphas were observed in two instances, suggesting potential deviations from market efficiency or study limitations.

Firstly, the recent subperiod from 2017 – 2023 revealed either statistically significant positive alphas for sustainable or long-short portfolios, or negative alphas for unsustainable portfolios. However, these findings were only partially robust to alternate portfolio weighting approaches or multi-factor regressions. This trend may reflect evolving regulations and growing consumer awareness or motivation, potentially driving momentum toward sustainable assets and overvaluing or correcting undervalued asset prices. However, this observation could also stem from limitations in the study, particularly in the selection of ESG data providers, as rating discrepancies among providers could significantly influence the findings. Furthermore, if indeed a deviation from market efficiency was observed, it is unlikely that alphas would persist over time. If this deviation stemmed from a price correction due to the previous lack of consideration of material ESG information, current prices would now reflect their fair value. Alternatively, if this deviation was driven by a demand hype around sustainable stocks, prices would now be overvalued, leading to lower future performance. Similar risk-adjusted returns over the 14-year period further support the notion that deviations would likely occur only for smaller subperiods, with a return to market efficiency probable over extended periods.

Secondly, a contradictory observation with statistically significant negative alphas surfaced when analysing value-weighted portfolios in the U.S. context. One explanation could be linked to reduced diversification and a specific bias in the ESGC score, particularly pronounced in the U.S. market. The largest companies received disproportionately lower ESGC scores due to heightened media scrutiny leading to increased controversies. This, coupled with the historical outperformance of the largest companies relative to the index, could have contributed to the negative alphas.

Despite numerous studies suggesting possible benefits of ESG during high volatility periods, the analysis of the COVID-19 pandemic yielded inconclusive or neutral findings. No statistically significant alphas were observed, thus failing to explain the previously mentioned positive alpha in the 2017 – 2023 subperiod. Risk-adjusted returns remained similar during the recovery but diverged during drawdown periods, although not to a significant extent. While the relationship was positive in the U.S., it was negative in Europe, suggesting a need for further exploration, especially given the relatively under-researched European context. A potential limitation lies in the methodology's failure to fully capture asymmetrical benefits like reduced downside risk and volatility. Future research directions could involve focusing solely on high-volatility periods to re-examine these findings or analysing the post-COVID-19 pandemic period to investigate the positive alpha observed in the recent subperiod.

Incorporating an additional ESGC weighting overlay onto traditional equal- or market value-weighted portfolios significantly enhanced the sustainability profile while maintaining similar risk-adjusted returns over extended periods and slightly higher risk-adjusted returns in the recent subperiod. Thus, the findings mirror those of the top and bottom ESGC score portfolios. Acknowledging the increased turnover associated with integrating ESGC weights, various levels of additional turnover-related fees were introduced. Results indicate that these extra fees were minimal, leading to overall similar net returns. In line with efficient markets, investors should expect to pay a slight premium due to slightly higher turnover-related fees. However, this study demonstrated that this does not significantly disadvantage sustainable investors. This underscores ESGC weighting as an intriguing approach to enhancing sustainability without screening out stocks or industries and potentially missing out on diversification opportunities.

The response to the research question indicates that integrating ESGC data into stock portfolio construction does not significantly affect risk-adjusted returns over extended periods. As a result, investors cannot leverage this information to improve their risk-adjusted returns, yet they need not be concerned about a decline in performance. Instead, they can employ it to increase the average ESGC score of their portfolio, for instance, through ESGC-weighted portfolios, thus boosting sustainability without compromising financial performance.

This thesis has substantial implications for investors and asset managers, highlighting two key findings. Firstly, investors can integrate sustainability into their investment strategies without sacrificing risk-adjusted returns. However, achieving this necessitates accessible and cost-effective integration methods, underscoring the significance of passive-managed ETFs that offer advanced approaches beyond basic negative screening. Secondly, asset managers may consider innovative products like passive managed ESG-tilted indexes, which could offer comparable risk-adjusted returns while boosting sustainability without excluding companies. This could serve as a unique selling point in a potentially saturated market landscape.

## **Acknowledgements**

I would like to dedicate this thesis to the memory of my beloved father, Manuel Neu, who graduated from this esteemed university in 1995 as an “Ingénieur commercial”. His academic achievements and the values he imparted have been a constant source of motivation for me.

I am deeply thankful to my supervisor, Professor Mathias Schmit, whose mentorship, guidance, and insightful feedback have been invaluable throughout the research process. His expertise and the impactful Banking and Asset Management course greatly influenced the direction of this master's thesis.

I would also like to thank CEBRIG, one of the oldest research centres at SBS-EM, and LSEG, the London Stock Exchange Group (formerly Refinitiv), for generously providing access to Datastream.

Special thanks go to my family and girlfriend for their unwavering support, understanding, and encouragement during this challenging academic endeavour. Their patience and willingness to assist have been instrumental in bringing this work to successful fruition.

## **Abstract**

This paper investigated the relationship between ESG scores and an investor's risk-adjusted financial performance. Refinitiv ESG combined scores were used to construct top, bottom, and ESG-tilted portfolios based on constituents from the S&P 500 and STOXX Europe 600 indices. Through quantitative analysis and multi-factor regression models, the study revealed that investors experienced statistically similar risk-adjusted returns over the 2010-2023 period, irrespective of adopting a sustainable or unsustainable investment approach. However, the recent subperiod from 2017 onwards showcased improved performance in sustainable portfolios, hinting at shifts in the landscape and heightened consumer awareness or motivation, possibly leading to recent deviations from the efficient market. Amidst the COVID-19 pandemic, no statistically significant differences were uncovered, underscoring the resilience of sustainable investment strategies during times of financial distress and failing to explain the recent shift in performance. Finally, this paper found that ESG-weighted portfolios provide risk-adjusted returns comparable to traditional portfolios while enhancing sustainability and maintaining diversification. While investors faced a slight premium for increased turnover to align portfolios with sustainability preferences, the findings indicated no statistically significant adverse effects on investors. Investors can seamlessly integrate sustainability criteria into their portfolios without fearing any adverse impact on risk-adjusted performance over the long term, although they should not anticipate superior performance outcomes. The key factor lies in adopting cost-effective and straightforward methods to introduce sustainability criteria, advocating for more advanced approaches beyond basic negative screening.

# Table of Contents

<b>List of Abbreviations.....</b>	<b>9</b>
<b>1 Introduction .....</b>	<b>10</b>
1.1 Motivation and Research Question .....	10
1.2 Delimitations .....	11
1.3 Thesis Structure .....	12
<b>2 Background .....</b>	<b>13</b>
2.1 Integration of Sustainability in Investment Strategies.....	13
2.2 ESG Scores and their Limitations .....	15
2.3 Regulatory Landscape.....	16
<b>3 Literature Review of Theories.....</b>	<b>19</b>
3.1 Efficient Market Hypothesis .....	19
3.2 Behavioural Finance .....	20
3.3 Modern Portfolio Theory .....	20
3.4 Capital Asset Pricing Model & Jensen's Alpha .....	21
3.5 Multi-factor Asset Pricing Models.....	23
3.5.1 Fama-French 3-Factor Model.....	23
3.5.2 Carhart 4-Factor Model.....	24
3.5.3 Fama-French 5-Factor Model.....	24
3.6 Theoretical Perspectives on ESG's Influence on Corporate Financial Performance .....	25
3.6.1 Negative ESG - Corporate Financial Performance Models.....	25
3.6.2 Positive ESG - Corporate Financial Performance Models.....	26
3.7 Theoretical Framework for Assessing the Impact of ESG on Investor Returns .....	27
<b>4 Literature Review of Empirical Studies .....</b>	<b>29</b>
4.1 Empirical Insights on the Relationship between ESG and Corporate Financial Performance.....	29
4.1.1 Corporate Social Performance and Corporate Financial Performance .....	29
4.1.2 Corporate Environmental Performance and Corporate Financial Performance .....	30
4.1.3 Corporate Governance and Corporate Financial Performance .....	31
4.1.4 Summary: ESG and Corporate Financial Performance.....	33
4.2 Empirical Investor Studies on the Relationship between ESG and Financial Performance.....	33
4.2.1 ESG Screening at Portfolio Level.....	34
4.2.2 Empirical Findings of Portfolio Studies.....	36
4.2.3 Conclusion & Reasons for Divergence.....	40
4.2.4 Insurance Effect during Economic Downturns .....	41
<b>5 Data &amp; Methodology .....</b>	<b>44</b>
5.1 Research Approach.....	44
5.2 Data.....	45
5.2.1 Stock Data and Studied Time Period .....	45
5.2.2 ESG Data – Refinitiv ESGC score .....	46
5.2.3 Integration and Refinement of Data Sources.....	49

5.2.4	Descriptive Statistics .....	50
5.3	<i>Methodology</i> .....	53
5.3.1	Portfolio Formation & Rebalancing.....	53
5.3.2	Performance Measurement .....	56
5.3.3	Multiple Linear Time Series Regression .....	58
<b>6</b>	<b>Empirical Analysis</b> .....	<b>60</b>
6.1	<i>United States of America</i> .....	60
6.1.1	Top/Bottom Portfolio Analysis (U.S.) .....	60
6.1.2	COVID-19 Period (U.S.) .....	63
6.1.3	ESG-weighted Portfolio Analysis (U.S.) .....	65
6.1.4	Robustness Tests (U.S.).....	68
6.2	<i>Europe</i> .....	69
6.2.1	Top/Bottom Portfolio Analysis (Europe) .....	69
6.2.2	COVID-19 Period (Europe) .....	71
6.2.3	ESG-weighted Portfolio Analysis (Europe).....	73
6.2.4	Robustness Tests (Europe) .....	75
<b>7</b>	<b>Discussion</b> .....	<b>77</b>
7.1	<i>ESG and Market Efficiency</i> .....	77
7.2	<i>Exploration of the COVID-19 Period</i> .....	80
7.3	<i>Examining the Viability of ESG-Weighted Portfolios</i> .....	82
7.4	<i>Limitations &amp; Suggestions for Further Research</i> .....	84
<b>8</b>	<b>Conclusion</b> .....	<b>87</b>
	<b>References</b> .....	<b>89</b>
	<b>Appendices</b> .....	<b>94</b>
	<i>Appendix 1: Statistical Tests</i> .....	94
	Appendix 1.1: Testing for Stationarity .....	94
	Appendix 1.2: Testing for Autocorrelation .....	95
	Appendix 1.3: Testing for Heteroscedasticity .....	96
	<i>Appendix 2: Robustness Analysis Regression Tables for the United States</i> .....	97
	Appendix 2.1: Value Weighted T20V and B20V Portfolios (U.S.) .....	97
	Appendix 2.2: Carhart 4-Factor Regression (U.S.) .....	98
	Appendix 2.3: Exclusion of Outliers (U.S.) .....	99
	<i>Appendix 3: Robustness Analysis Regression Tables for Europe</i> .....	101
	Appendix 3.1: Value Weighted T20V and B20V Portfolios (Europe).....	101
	Appendix 3.2: Carhart 4-Factor Regression (Europe).....	102
	Appendix 3.3: Exclusion of Outliers (Europe).....	103
	<i>Appendix 4: Impact of ESGC Weighting on Sustainability</i> .....	105
	<i>Appendix 5: Annual Turnover Rates</i> .....	106
	<i>Appendix 6: ESGC Score Bias – Example of S&amp;P 500</i> .....	108
	<i>Appendix 7: ESGC Score Outliers</i> .....	110

# List of Abbreviations

Abbreviation	Definition
AUM	Assets under Management
B20E	Portfolio: Bottom 20% of companies by ESGC score (equal-weighted)
B20V	Portfolio: Bottom 20% of companies by ESGC score (market value-weighted)
CEP	Corporate Environmental Performance
CFP	Corporate Financial Performance
CG	Corporate Governance
CMA	Risk factor: Conservative minus Aggressive
CSP	Corporate Social Performance
CSR	Corporate Social Responsibility
CSRD	Corporate Sustainability Reporting Directive
EMH	Efficient Market Hypothesis
ESG	Environmental Social Governance
ESGC	Specific ESG score provided by Refinitiv: ESG Combined score
ESGE	Portfolio: ESGC-weighted market portfolio
ESGV	Portfolio: ESGC- and market value-weighted market portfolio
ESRS	European Sustainability Reporting Standards
ETF	Exchange Traded Fund
EU	European Union
HAC	Heteroscedasticity and autocorrelation consistent
HML	Risk factor: High minus Low
MKTE	Portfolio: equal-weighted market portfolio
MKTV	Portfolio: market value-weighted market portfolio
MPT	Modern Portfolio Theory
NFRD	Non-Financial Reporting Directive
NRBV	Natural Resource-Based View
OLS	Ordinary Least Squares
RMW	Risk factor: Robust minus Weak
SFDR	Sustainable Finance Disclosure Regulation
SIC	Stakeholder Influence Capital
SMB	Risk factor: Small minus Big
SRI	Socially Responsible Investment
T20E	Portfolio: Top 20% of companies by ESGC score (equal-weighted)
T20V	Portfolio: Top 20% of companies by ESGC score (market value-weighted)
U.K.	United Kingdom
U.S.	United States of America
WML	Risk factor: Winner minus Loser

# 1 Introduction

## 1.1 Motivation and Research Question

The financial landscape has undergone a strong transformation in recent years, marked by an increasing focus on critical global issues such as climate change, biodiversity, human rights, and social inequality. These concerns have ascended to the forefront of public and political attention, leading to a heightened awareness within the finance industry. The Principles for Responsible Investment (PRI), a United Nations-backed initiative, defines responsible investing as an approach that seeks to integrate Environmental, Social, and Governance (ESG) factors into investment decisions<sup>1</sup>. This integration aims to enhance risk management strategies and generate sustainable, long-term returns. Incorporating ESG principles into investment practices has shifted from being a niche strategy to a mainstream consideration for market participants, underscoring a broader recognition within the finance community that factors beyond traditional financial metrics play a pivotal role in shaping investment decisions and outcomes.

With the increasing mainstream adoption of ESG principles and the implementation of regulations regarding ESG and financial instruments, numerous studies have sought to evaluate the impact of ESG on financial performance. However, findings have not consistently aligned. The variability in results is often linked to factors such as the type of data considered, the chosen geographical focus, the methodology employed, the periods studied, and more. In general, two major types of studies can be distinguished. First, corporate-focused studies, also known as nonportfolio studies, aim to assess the impact on companies integrating ESG into their principles. This assessment frequently incorporates metrics like return on equity, return on assets, or firm market value. Second, investor-focused studies, also known as portfolio studies, aim to evaluate the impact on investors when integrating ESG principles into their investment decisions. These studies may examine the direct relationship between ESG and financial performance by comparing portfolio return differences to benchmarks. While corporate-focused studies typically reveal a positive relationship, investor-focused studies often present mixed or neutral findings, leaving ample room for further research in this domain (Friede et al., 2015; Whelan et al., 2021).

Critics of sustainable investing argue that ESG-based screening narrows the investment universe, thereby reducing diversification and consequently lowering the risk-adjusted return on investment (Markowitz, 1952). Conversely, advocates contend that screening for high-ESG stocks results in a limited investment universe comprising superior companies that outperform the market, mitigating the adverse impact of reduced diversification (Barnett & Salomon, 2006). The impact of sustainable investing on financial returns holds substantial implications for both investors and corporations.

---

<sup>1</sup> Principles for Responsible Investment. (n.d.). *What is responsible investment?* Retrieved March 10, 2024, from <https://www.unpri.org/introductory-guides-to-responsible-investment/what-is-responsible-investment/4780.article>

Sustainability factors may emerge as a crucial consideration in investors' decisions regarding asset allocation if they enhance financial returns rather than diminish them.

In this context, the primary goal of this study is to examine the relationship between ESG characteristics and risk-adjusted financial performance in the portfolio context. The aim is to explore whether ESG considerations positively influence financial performance and, if so, to offer insights into how investors can craft investment strategies to leverage potential market inefficiencies. Understanding the potential link between ESG ratings and financial performance has become pivotal in asset managers' and retail investors' stock selection and portfolio construction. The central inquiry revolves around the feasibility of achieving financial success while simultaneously contributing to societal well-being. A positive relationship could attract even the most sceptical investors to ESG investing, while a neutral relationship would validate the efficient market hypothesis. In either case, the findings would underscore the notion that financial success and positive contributions to society are not mutually exclusive. Therefore, this study seeks to address the following research question:

**" To what extent does the integration of ESG data in the formation of stock portfolios influence risk-adjusted returns, and can investors devise strategies to leverage this information effectively? "**

The central research question guides the analysis. However, this study explores multiple dimensions, shedding light on variations in ESG performance across regions, specifically focusing on the United States of America and Europe over an extended period. Furthermore, it delves deeper into economic downturns, examining the performance of ESG portfolios during crisis periods with an analysis of the COVID-19 pandemic, encompassing both the drawdown and recovery periods. Additionally, diverse asset allocation strategies are analysed, including frequently studied long-short strategies and more innovative ESG-tilted portfolios. Finally, all these insights are provided from an investor-centric perspective, assessing the financial impact for an investor aligning their portfolio with ESG considerations and considering overlooked aspects such as turnover fees.

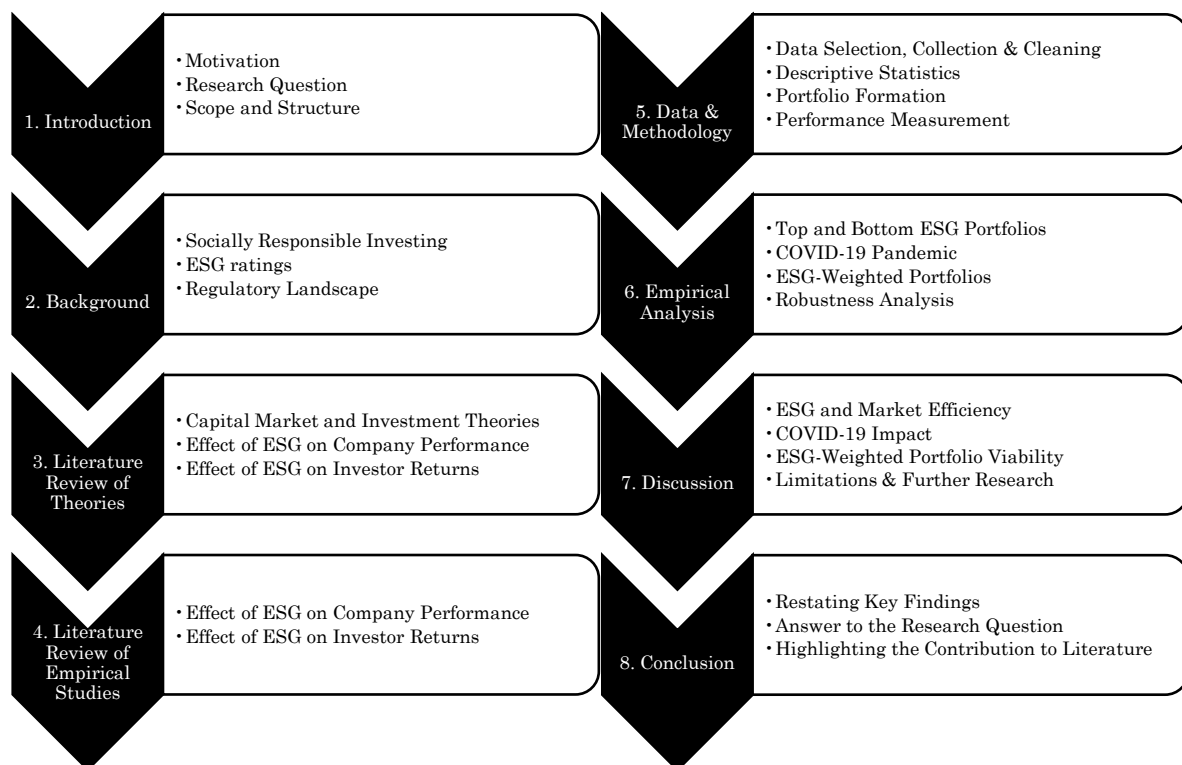
## **1.2 Delimitations**

Certain limitations are imposed on the study's scope to enhance its depth. Primarily, the focus is on overall ESG ratings, refraining from a detailed exploration of individual pillars or category variables. This deliberate choice aligns with the study's goal of catering to regular investors seeking practical insights into integrating ESG criteria rather than identifying specific past anomalies susceptible to overfitting. The approach remains general, designed to examine ratings that can feasibly guide investors in incorporating sustainability considerations into their portfolios. Furthermore, this study refrains from making industry-specific analyses, recognising that industry-specific outcomes may be influenced by unique circumstances over time, potentially diverging from the broader market context. Moreover, acknowledging investors' typical pursuit of portfolio

diversification in terms of both the number of companies and industry allocations, the assessment centres on the overarching implications of integrating ESG considerations into a diversified portfolio. The evaluation prioritises the investor’s viewpoint and focuses solely on market-based financial performance measures, steering clear of accounting-based metrics such as operating profitability. Despite the inherent correlation between these measures, the aim is to investigate whether investors can achieve higher or lower risk-adjusted returns by integrating general ESG criteria into their investment decisions.

### 1.3 Thesis Structure

The thesis is organised into eight chapters to address the primary research question and analyse various aspects of the subject matter. Chapter 1 serves as an introduction, providing insights into the study’s motivation, research question, and scope. Chapter 2 offers background information on ESG investing, covering the conceptual framework, ESG ratings, and the regulatory landscape. The literature review is divided into two segments: Chapter 3 outlines the theoretical concepts related to financial markets, ESG incorporation for companies, and ESG investing, while Chapter 4 explores empirical findings on these subjects. Chapter 5 details the data and methodology employed. The empirical results are reported in Chapter 6, including an in-depth examination of the European and U.S. markets, and subsequently discussed in Chapter 7. The conclusion of the thesis is provided in Chapter 8.



## 2 Background

### 2.1 Integration of Sustainability in Investment Strategies

Integrating sustainability considerations into investment strategies has emerged as a pivotal topic in the contemporary landscape of financial markets. In response to the evolving dynamics of environmental, social, and governance (ESG) factors, investors and financial institutions increasingly recognise the significance of incorporating sustainability criteria into their decision-making processes. This shift reflects a broader acknowledgement of the interconnectedness between financial performance and non-financial, sustainable practices. As global awareness grows regarding climate change, social equity, and responsible corporate governance, the investment community is driven to explore ways to align financial objectives with positive environmental and societal impact.

This trend is evident globally, with sustainable Assets under Management (AUM) experiencing consistent growth year after year. The 2022 Global Sustainable Investment Alliance report revealed a significant 20% increase in sustainable AUM from 2020 to 2022<sup>2</sup>. The cumulative global investment in sustainable assets reached an impressive \$30.3 trillion, comprising approximately 25% of the total AUM. These figures underscore the sustained and widespread momentum of sustainable investing across diverse financial markets (GSIA, 2022).

Over time, various sustainable investment strategies have emerged, with Eurosif (2018) categorising them into seven fundamentally different approaches:

- Best in class: This approach involves selecting or weighting investments based on ESG criteria, specifically focusing on the best-performing or most improved companies or assets within a defined investment universe.
- Engagement & voting: This strategy involves engaging in long-term activities such as voting shares and actively participating in discussions with companies on ESG matters to influence behaviour or enhance disclosure.
- ESG integration: Asset managers systematically incorporate ESG risks and opportunities into traditional financial analysis, considering both ESG and financial factors in the mainstream analysis of investments. The integration process emphasises the potential impact of ESG issues on company financials, influencing investment decisions.
- Exclusions: This approach systematically excludes specific investments, such as companies, sectors, or countries, from the investible universe based on criteria like weapons, pornography, tobacco, and animal testing.
- Impact investing: This strategy involves allocating capital to companies, organisations, and funds to generate social and environmental impact alongside a

---

<sup>2</sup> This trend does not encompass the U.S. market, given a methodology change in the U.S. data that hinders a consistent comparison across regions and time

financial return. It distinguishes itself from philanthropy by retaining ownership of the asset and expecting a positive financial return.

- Norms-based screening: The process of evaluating investments for adherence to international standards and norms, typically established by international entities like the United Nations.
- Sustainability-themed: This approach entails investing in assets or funds that focus on specific ESG issues, contributing to the resolution of challenges like climate change, eco-efficiency, and health.

Krosinsky (2014) emphasises the need for a positive, future-oriented, and opportunity-driven paradigm that prioritises social and environmental considerations to add value for all investors and effectively helps solve sustainability issues. He argues that divesting from certain industries, such as through basic exclusionary screens, falls short of making sustainable investment meaningful. According to Folqué et al. (2021), ESG integration is a far more advanced method to mitigate ESG risks than conventional negative filters. They contend that negative filters, one of the oldest and most basic sustainable investment practices, were initially used in the absence of data, but they are not the most suitable strategy for reducing ESG and carbon risk in investment portfolios.

Experiencing a compound annual growth rate of approximately 25% since 2016, ESG integration has emerged as the second-fastest-growing investment strategy, surpassing the exclusion strategy since 2020 to become the largest in terms of AUM (GSIA, 2020). Consequently, given its substantial market share and rapid growth, there is a need to examine the relationship between ESG metrics and financial performance. It is believed that considering financially material ESG metrics leads companies to perform better in the long run. Companies prioritising ESG criteria or implementing robust ESG risk management strategies may be better equipped to handle non-financial risks associated with climate change, data fraud, theft, and social issues (Antoncic, 2019).

In the initial phases of responsible investing, individuals were primarily motivated by personal values such as ethics and environmental considerations rather than financial performance (Renneboog et al., 2011). However, as ESG gained widespread acceptance, financial performance emerged as another significant factor influencing sustainable investing. A survey conducted by Amel-Zadeh & Serafeim (2018) revealed that financial considerations have become the primary motivation for utilising ESG information, surpassing ethical concerns. The belief is that ESG information is considered material to investment performance. Giglio et al. (2023) found in a large panel survey of retail investors' preferences and beliefs on ESG, that significant ESG holdings are kept exclusively among investors anticipating outperformance in the market, even among those who cited ethical or climate hedging reasons as their foremost motivations for ESG investments.

## 2.2 ESG Scores and their Limitations

Commercially developed, third-party ESG ratings have emerged to meet various stakeholders' growing demand for ESG information. Asset owners seek ESG data to understand their investments' environmental and social impact. Institutional investors aim to create ESG-focused products for their clients and assess the potential financial implications of societal and environmental factors on companies' performance. Companies themselves use ESG scores to showcase their sustainability efforts, acting as both suppliers and consumers of ESG information. Regulators are interested in ESG data to evaluate investment managers' claims about ESG integration in their processes. Other stakeholders, such as local governments, consultants, and university students, are keen to access ESG scores for various purposes. This increasing interest in ESG ratings highlights their significance in the investment landscape and the broader business community (Larcker et al., 2022). The rating agencies have successfully promoted transparency among companies regarding their sustainability efforts by consolidating this information into a single score. This simplified numerical representation can be readily used by all stakeholders involved.

ESMA (2022a) comprehensively defines ESG rating as “an opinion regarding an entity, issuer, or debt security’s impact on or exposure to ESG factors, alignment with international climatic agreements or sustainability characteristics, issued using a defined ranking system of rating categories”. These ESG ratings encompass two primary perspectives, each with distinct focuses (Larcker et al., 2022). The first viewpoint centres on evaluating a company's influence on its stakeholders, where adopting responsible business practices and avoiding harmful activities can lead to higher scores. These are known as "ESG impact ratings" as they emphasise the company's societal and environmental impact. The second perspective involves assessing how societal and environmental factors affect the company. The ESG framework identifies risk factors that the company can address through strategic planning, targeted investments, or operational adjustments. These ratings are known as "ESG risk ratings".

ESG ratings continue to encounter criticism. The first criticism can be attributed to the presence of numerous providers in the market, each using distinct methodologies with varying levels of transparency in their calculations. As a result, ESG scores from different providers show only moderate correlations with one another. A study by Berg et al. (2022) compared ratings from five different providers and revealed a significant divergence in ESG ratings. The correlation between ESG scores from Sustainalytics and Refinitiv was just 67%. Substantial disparities were evident, particularly in categories related to corporate governance, human rights, indigenous rights, health, safety, and other aspects. Other providers showed even worse correlations, ranging between 37% to 71%. Dimson et al. (2020) and Zumente & Lāce (2021) came to similar findings. It is crucial to consider this finding, given that prior studies exploring the relationship between ESG and financial performance employed various ESG data providers. Consequently, the conclusions drawn from such studies may not be directly comparable to other studies using data from different ESG rating providers. The fact that many ESG rating agencies do not disclose to the public their complete methodological guides to how their ESG

ratings are calculated aggravates the problem because the lack of transparency reduces the possibility of correctly comparing ratings. Moreover, Berg et al. (2022) suggest that the divergence in ESG ratings is not merely due to differing definitions but is rooted in fundamental disagreements about the underlying data. The ideal foundation for ESG ratings should be objective observations that can be reliably ascertained and agreed upon across the industry.

The second critique highlights that utilising various information sources, including companies' voluntary disclosures, surveys, third-party data, and unstructured company data, leads to inconsistencies in ESG ratings. Due to the lack of standardisation and structure in these disclosures, it becomes challenging to make accurate comparisons and interpretations (Sipiczki, 2022). Christensen et al. (2022) discovered that increased ESG disclosures lead to greater disagreement among ESG ratings. This indicates that rating companies struggle to interpret the disclosed information due to its subjective nature. Furthermore, the reliance on self-reported data raises concerns about the reliability and quality of such information. Del Giudice & Rigamonti (2020) discovered that third-party auditing of sustainability reporting can improve the reliability of ESG scores. This auditing process adds an assurance effect to the quality of a company's ESG information, suggesting that audited ESG reports are more trustworthy and accurate than unaudited ones. Moreover, the "State of Play in Reporting and Assurance of Sustainability Information" reported an increasing trend in ESG assurance provided on companies' reporting. In 2020, 58% of companies had ESG assurance (IFAC, 2022).

Apart from the two main criticisms mentioned earlier, ESG ratings are influenced by biases related to company size and geographical location, as noted by Sipiczki (2022). Company size bias indicates that, on average, larger firms tend to have higher ESG scores not necessarily due to better ESG behaviour or risk management but primarily because of their increased disclosure of ESG data (Drempetic et al., 2020). Larger companies typically allocate higher budgets for disclosing ESG data and face more significant pressure to be transparent. For instance, a study by Tamimi & Sebastianelli (2017) using Bloomberg ESG scores found that large-cap companies within the S&P 500 have significantly higher ESG disclosure scores than mid-cap companies. As ESG scores are based on disclosed information, this increased disclosure rate leads to higher ESG scores. Similar to company size bias, geographical bias is due to different levels of disclosures in different regions. For example, ESG scores in Europe are considerably higher than those in the United States, which can be attributed to Europe's more developed sustainable reporting landscape compared to the absence of a mandatory reporting framework in the United States (Sipiczki, 2022).

## **2.3 Regulatory Landscape**

The divergence in ESG scores from different providers can be attributed to the lack of regulation. In recent years, the European Union (EU) has taken on a leading role in regulating sustainable disclosures by companies listed or operating within its borders. In contrast, other countries have been slower in implementing such regulations. Europe's

proactive stance in setting regulations may stimulate other nations to follow suit or adopt the European standard.

With the introduction of the new EU ESG regulations, companies operating in international markets may face growing challenges when reporting using different standards. As the EU regulations get implemented across Member States, the expenses associated with compliance might discourage companies from adhering to their usual reporting standards. This could pave the way for the EU ESG standards to become a universal benchmark for sustainability reporting worldwide, leading to a homogenisation of the reporting standards (Redondo Alamillos & de Mariz, 2022).

Standardising reporting practices represents an initial step towards reducing the divergence in ESG scores. Dumrose et al. (2022) discovered that the EU Taxonomy and the requirement to disclose environmental data could reduce measurement divergence in E ratings. This reduction in divergence can help mitigate greenwashing risks within the financial system. The EU Taxonomy is a classification tool by the European Commission that determines whether an activity is deemed environmentally sustainable. As the European Commission currently envisions extending the EU Taxonomy to social objectives (Platform on Sustainable Finance, 2022), this standardisation could further reduce measurement divergence across the global ESG ratings.

Europe has witnessed a growing trend of regulatory developments, starting with the implementation of the Non-Financial Reporting Directive, NFRD (Directive 2014/95) in 2014. This directive mandated listed companies and firms above certain thresholds of capital, sales, and employees to include a sustainability statement alongside their financial reports. The NFRD was replaced by the Corporate Sustainability Reporting Directive, CSRD (Directive 2022/2464) which went into force in January 2023. The CSRD expands the number of companies subject to EU sustainability reporting requirements and reinforces the regulatory framework. On 31<sup>st</sup> of July 2023, the Commission adopted the European Sustainability Reporting Standards (ESRS) for use by all the companies subject to the CSRD<sup>3</sup>. As explained by Frecautan & Nita (2022), “an ESG standard is very specific outlining criteria on how the data needs to be collected”, which makes frameworks more practical and marks an important step towards improving the reliability and comparability of disclosed data. Moreover, the CSRD requires assurance on the sustainability information that companies report. Next to the NFRD and CSRD, the Sustainable Finance Disclosure Regulation, SFDR (Regulation 2019/2088) applies as of March 2021 and sets a comprehensive reporting framework to improve transparency and prevent greenwashing by setting disclosure requirements for financial products that claim to have ‘sustainable investment’ as their objective (Article 9 funds) and financial

---

<sup>3</sup> Directorate-General for Financial Stability, Financial Services and Capital Markets Union. (2023, July 31). The Commission adopts the European Sustainability Reporting Standards. *European Commission*. Retrieved March 10, 2024, from [https://finance.ec.europa.eu/news/commission-adopts-european-sustainability-reporting-standards-2023-07-31\\_en](https://finance.ec.europa.eu/news/commission-adopts-european-sustainability-reporting-standards-2023-07-31_en)

products that claim to be promoting social or environmental characteristics (Article 8 funds).

This summary highlights the European Commission's efforts in a short timeframe to develop directives and regulations aimed at expanding the applicability of these rules to more companies and improving the quality of their disclosures. These measures directly address the earlier identified problems, such as ESG rating divergence and inconsistencies. The introduction of mandatory ESG disclosures within a standardised framework and data assurance fosters uniformity in disclosure, enhancing comparability and facilitating interpretation for rating providers. As the CSRD has recently come into force, and companies are only required to disclose information from 2024 onwards, there is currently limited existing literature on the extent to which these directives and standards have brought about changes. However, it can be anticipated that future studies will emerge to assess the extent to which the CSRD and ESRS have contributed to enhancing the quality of ratings.

Introducing mandatory ESG disclosure offers various other benefits, such as enhancing overall ESG performance for companies subject to these requirements, as seen in the case of NFRD (Aluchna et al., 2022) and SFDR implementation (Becker et al., 2022). It also favours a company's corporate information environment, resulting in positive capital market outcomes. This includes enhancements in analysts' earnings forecasts and increased availability and improved quality of ESG reporting. Consequently, this helps mitigate the risk linked to negative ESG incidents and subsequent stock price crashes (Krueger et al., 2021).

The absence of a mandatory sustainable reporting framework in the United States starkly contrasts the European Union, making it particularly intriguing to compare the relationship between ESG and financial performance between the American and European markets. However, there are indications of progress in the United States, as the Securities and Exchange Commission is currently working on the "Enhancement and Standardization of Climate Related Disclosures" (SEC 33-11042/34-94478) which would require certain companies to report on climate-related risks and their impact on business, strategy, and outlook. While the U.S. is yet to adopt a comprehensive reporting framework, this initiative signals a step towards increased climate-related disclosure obligations for certain companies.

### 3 Literature Review of Theories

The foundation of capital market theory lies in the idea of an efficient market, where participants trade financial securities intending to maximise risk-adjusted returns. This theory operates on the belief that stock prices fully reflect all available information, and investors cannot consistently outperform the market, emphasising the trade-off between risk and potential rewards in the investment process.

#### 3.1 Efficient Market Hypothesis

The efficiency of capital markets is measured by their ability to determine security prices accurately. This involves two crucial aspects: the speed at which prices adjust and the magnitude and direction of the adjustment. The Efficient Market Hypothesis posits that security prices represent the most accurate estimation of their investment value. Since capital markets automatically incorporate new information into prices, they promptly adjust to reflect the latest developments. As a result, investors cannot systematically earn returns above the market average without engaging in speculative activities that carry significant risks. The underlying economic rationale suggests that when information indicates that a stock is undervalued and offers a profit potential, investors rush to buy the stock, leading to an immediate increase in its price to a fair value, where only normal returns can be anticipated (Bodie et al., 2018).

As outlined by Fama (1970), the efficient market hypothesis can be divided into three forms: weak, semi-strong, and strong.

In the weak form, current security prices are believed to incorporate all information contained in past prices. Consequently, studying past performance will not provide any predictive insight into future security prices. This form posits that only new information, not historical data, can prompt changes in security prices.

Moving to the semi-strong form, the hypothesis contends that security prices not only reflect all past prices but also encompass all publicly available information. Consequently, analysing publicly available information will not allow investors to achieve abnormal or excess returns, as security prices swiftly adjust to new information.

Lastly, the strong form asserts that security prices reflect not only publicly available information but also all privately held information. In this theory, no investor can achieve abnormal returns, even with access to insider information.

Typically, the prevailing view is that markets exhibit semi-strong efficiency.

In real-world scenarios, some instances raise questions about the validity of the efficient market hypothesis, often referred to as market anomalies. A market anomaly arises when the price of a stock diverges from the behaviour predicted by a model. Malkiel (2003) highlights several well-known anomalies, such as the size effect, which shows that in many markets, small firms outperform larger ones and generate higher risk-adjusted returns over the long term. Another anomaly is the P/E effect, where firms with low price-to-earnings (P/E) ratios tend to yield abnormal returns compared to those with high P/E ratios. Furthermore, there are seasonal and day-of-the-week patterns, such as the

January effect, wherein stocks, especially small-sized ones, exhibit unusually high returns during January. These anomalies challenge the assumptions of the efficient market theory and reveal complexities in stock market behaviour.

Two contrasting views have emerged to explain anomalies in the financial markets. On one side, Fama & French (1998) argue that anomalies may arise due to asset pricing theories or simply result from random chance. On the other hand, Kahneman & Tversky (1979) propose behavioural approaches, suggesting that anomalies can be attributed to investors' cognitive limitations, leading to irrational investment decisions. Fama & French (2007) contend that conventional asset pricing models make unrealistic assumptions and overlook the impact of these preferences. They posit that investors' tastes and preferences can affect asset prices and potentially impact investment performance.

### **3.2 Behavioural Finance**

In contrast to classical theories that assume investors to be rational, behavioural finance perceives investors as inherently irrational. Investors are susceptible to various cognitive errors when processing information, often influenced by biases resulting in suboptimal decision-making. According to Bodie et al. (2018), these biases include forecasting errors, where individuals succumb to memory bias, assigning excessive weight to recent experiences and making overly extreme forecasts despite inherent uncertainty; overconfidence, leading investors to believe they can outperform the market; conservatism, where investors update their beliefs slowly in response to new evidence, leading to underreaction to news; and representativeness bias, where investors may treat small samples as if they are representative of the entire population. In addition to challenges related to information processing, investors grapple with additional biases, including emotions and feelings that consumers might associate with a potential purchase or investors with a stock. These emotional factors can subsequently impact stock prices and their overall profitability.

Certainly, there is a consensus that investors' decisions are not always optimal. However, the degree to which limited rationality influences asset pricing is a matter of ongoing debate, as well as whether one can capitalise on such mispricing. Malkiel (2003) contends that the stock market is not always perfectly efficient, acknowledging that investors can occasionally make mistakes, leading to pricing irregularities and predictable patterns in stock returns. While sporadic deviations from efficiency may occur, the prevailing belief persists that the stock market is generally efficient in processing information, and any irregularities are unlikely to persist.

### **3.3 Modern Portfolio Theory**

Introduced by Markowitz (1952), Modern Portfolio Theory (also known as mean-variance analysis) underscores the significance of diversification in reducing risk for investors. The idea is that the variance in stock returns is exposed to both systematic risk

and the firm-specific risk inherent in a particular investment due to its unique characteristics. However, by combining uncorrelated risky assets in a portfolio, investors can effectively diversify away the firm-specific risk, resulting in a reduction of the overall risk profile without sacrificing expected returns. Systematic risk, also known as market risk, is inherent in all stocks and persists despite diversification efforts.

Modern Portfolio Theory (MPT) is grounded in the assumption that investors, being risk-averse, face a trade-off between risk and return in the securities market. This trade-off implies that investors anticipate a premium when confronted with heightened volatility, establishing a direct relationship between higher systematic risk and increased expected returns. Various ratios, like the Sharpe ratio, assess the risk-return relationship of a portfolio, enabling a direct comparison of stock or portfolio performance.

MPT aims to construct portfolios that achieve either the highest expected return for a specified level of risk or the lowest risk for a desired rate of return. Central to this theory is the concept of mean-variance efficiency, determining the optimal trade-off between risk and return. A portfolio is considered mean-variance efficient when no alternative portfolio can offer a higher expected return for the same level of risk. As assumed in MPT, rational investors are mean-variance optimisers, consistently choosing portfolios lying on the efficient frontier. The efficient frontier represents the set of portfolios providing the maximum expected return for every level of portfolio risk (Bodie et al., 2018).

Introducing constraints on investment choices, such as maximum exposures to certain assets or turnover limits, will restrict the efficient frontier. Consequently, the expected return for a given level of risk may be reduced compared to a situation with no constraints. The constraints limit the range of possible portfolios, leading to a trade-off where investors may need to sacrifice some expected return to adhere to the imposed restrictions while managing risk (Fabozzi et al., 2002). Therefore, rational investors seeking to optimise mean-variance portfolios would typically prefer a portfolio located on the unconstrained efficient frontier, as it is more likely to offer a more favourable risk-return trade-off.

### **3.4 Capital Asset Pricing Model & Jensen's Alpha**

The Capital Asset Pricing Model (CAPM) is a single-factor model that establishes a relationship between the systematic risk and the expected return of risky assets. Developed in the 1960s by Sharpe (1964), Lintner (1965), and Mossin (1966), the CAPM is built on several key assumptions. Firstly, it assumes that investors are rational mean-variance optimisers who seek to maximise returns for a given level of risk. Secondly, it assumes that all investors share homogeneous market expectations, implying that all relevant information is publicly available and known to investors. Lastly, the model assumes that investors can access risk-free borrowing and lending, enabling them to take short positions on traded securities.

The CAPM is based on the Modern Portfolio Theory principle that a company's idiosyncratic risk can be adequately reduced through diversification, leaving it only affected by a single risk factor, the systematic risk. This non-diversifiable risk is quantified using the beta value, which is calculated as the covariance between a stock's return and the overall market's return, divided by the variance of the market return. In essence, the beta value helps gauge the extent to which a stock's returns move in relation to market movements, capturing the systematic risk component of the asset.

$$\beta_i = \frac{Cov(R_{it}, R_{mt})}{\sigma_m^2} \quad (1)$$

Where  $Cov(R_{it}, R_{mt})$  is the covariance between asset  $i$  and the market, and  $\sigma_m^2$  is the variance of the return on the market portfolio. The beta coefficient ( $\beta_i$ ) quantifies how much asset  $i$  covaries with the market, and it is directly related to the asset's expected return. The market beta is equal to one by definition, representing the average risk of the market. Assets with beta values above one fluctuate more than the market and, therefore, are exposed to higher systematic risk. Consequently, these assets are expected to yield higher returns to compensate for the increased risk.

The general CAPM was first expressed by Sharpe (1964):

$$E[R_{it}] = R_{ft} + \beta_i(R_{mt} - R_{ft}) \quad (2)$$

Where  $E[R_{it}]$  is the expected return of asset  $i$  in period  $t$ ,  $R_{ft}$  is the risk-free rate in period  $t$ ,  $\beta_i$  measures the systematic risk of asset  $i$  compared to the market portfolio, and  $R_{mt}$  is the return on the market portfolio.

In practice, the addition of a riskless asset to the efficient frontier of modern portfolio theory creates the Capital Market Line, represented by the tangent line to the efficient frontier. This inclusion expands investors' investment opportunities, enabling them to borrow or lend at the risk-free rate and combine these investments with the market portfolio. This allows investors to implement strategies that offer higher returns for the same level of risk or lower risk for the same level of return compared to the original efficient frontier in portfolio theory (Bodie et al., 2018).

Jensen's alpha, formulated by Jensen in 1968, serves as a metric for risk-adjusted returns. It illustrates the extent to which actual stock returns differ from the expected ones, commonly termed "abnormal returns". Jensen's alpha discerns potential underperformance or outperformance by indicating how a particular asset's performance compares to the projection outlined by the CAPM (Jensen, 1968). Thus, Jensen's alpha encapsulates the unexplained return beyond the control variables employed in the model. This abnormal return will be simply referred to as "alpha" throughout this thesis.

In the context of the CAPM, the equation is expressed as follows:

$$R_{it} - R_{ft} = \alpha_i + \beta_i(R_{mt} - R_{ft}) + \epsilon_{it} \quad (3)$$

Where:

$R_{it}$	return of asset i in period t
$R_{mt}$	return of the market portfolio in period t
$R_{ft}$	risk-free rate in period t
$\beta_i$	assets' i sensitivity to the market
$\alpha_i$	assets' i Jensen's alpha
$\epsilon_{it}$	assets' i error term in period t

### 3.5 Multi-factor Asset Pricing Models

The CAPM exhibits limitations attributed to well-known market anomalies. Depending solely on market beta, it struggles to sufficiently explain the fluctuations in stock returns. This underscores the rationale for employing more sophisticated multi-factor models, such as the Fama & French (1993) three-factor model, the Carhart (1997) four-factor model, or even the Fama & French (2015) five-factor model, which incorporate additional risk factors. This thesis uses alpha derived from a multi-factor model to assess the abnormal returns attained by the constructed ESG portfolios. To avoid repetition, only new factors will be presented for each model, without revisiting previously discussed ones.

#### 3.5.1 Fama-French 3-Factor Model

In response to empirical anomalies, particularly the tendency for small-cap and value stocks to outperform compared to CAPM predictions, Fama and French (1993) introduced their 3-factor model, which accounts for these market irregularities by incorporating two additional risk factors:

**The Size Factor (SMB - Small Minus Big):** This factor captures the historical tendency of small-cap stocks to outperform large-cap stocks. It assesses the return difference between the 50% smallest stocks and the 50% largest stocks by market capitalisation.

**The Value Factor (HML - High Minus Low):** This factor reflects the historical outperformance of value stocks (those with high book-to-market ratios) compared to growth stocks (those with low book-to-market ratios). It quantifies the return difference between the 30% cheapest stocks by price-to-book (P/B) ratio and the 30% most expensive stocks by P/B ratio.

The following equation expresses the Fama-French three-factor model, including alpha:

$$R_{it} - R_{ft} = \alpha_i + \beta_1(R_{Mt} - R_{ft}) + \beta_2SMB_t + \beta_3HML_t + \epsilon_{it} \quad (4)$$

This model proves highly effective, explaining over 90% of returns for diversified portfolios - significantly surpassing the explanatory power of the CAPM, which typically accounts for approximately 70% of the variation. Independent studies by Bello (2008), Blanco (2012), Rehnby (2016), and others validated its effectiveness, confirming its superior fit to empirical data.

### 3.5.2 Carhart 4-Factor Model

Following the development of the Fama and French 3-factor model, several extensions were introduced to enhance its explanatory power. A notable advancement in this direction was the 4-factor model proposed by Carhart (1997). Carhart's model builds upon the foundation of the Fama & French (1993) model by introducing a fourth factor: the momentum factor (**WML – Winner Minus Loser**). Its strategy is based on the idea that assets that have performed well recently will continue to perform well, while those that have performed poorly will continue to perform poorly.

The momentum factor in Carhart's model is designed to capture the historical performance of securities based on their recent price trends. It compares the returns on stocks that have shown exceptional performance over the past year to those that did not. WML is calculated as the average return on securities with the 50% highest 11-month returns (lagged by one month) minus the average return on securities with the 50% lowest corresponding returns.

The Carhart four-factor model, encompassing alpha, is represented by the following equation:

$$R_{it} - R_{ft} = \alpha_i + \beta_1(R_{Mt} - R_{ft}) + \beta_2SMB_t + \beta_3HML_t + \beta_4WML_t + \varepsilon_{it} \quad (5)$$

This addition to the Fama-French Three-Factor Model enhances its explanatory power from 90% to 95%. Moreover, other independent studies, such as those conducted by Bello (2008) and Rehnby (2016), have also found that this model offers superior explanatory power.

### 3.5.3 Fama-French 5-Factor Model

Another renowned extension is the Fama & French (2015) 5-factor model, which also explains approximately 95% of portfolio returns. Instead of incorporating the momentum factor, this model integrates two additional factors: Profitability and Investment. These factors were selected based on empirical findings indicating that stocks with robust operating profitability tend to outperform those with weak operating profitability (**RMW – Robust Minus Weak**), and stocks from companies with conservative investment practices outperform those with aggressive investment strategies (**CMA – Conservative Minus Aggressive**). The RMW factor represents the return of stocks with the 30% highest operating profitability minus the return of stocks

with the 30% lowest or negative operating profitability<sup>4</sup>. The CMA factor measures the return of stocks from companies needing minimal ongoing capital investment to maintain and grow their business against those needing significant ongoing capital investment, determined by the difference between the lowest 30% and highest 30% investment ratios<sup>5</sup>.

The following equation expresses the Fama-French 5-factor model, including alpha:

$$R_{it} - R_{ft} = \alpha_i + \beta_1(R_{Mt} - R_{ft}) + \beta_2SMB_t + \beta_3HML_t + \beta_4RMW_t + \beta_5CMA_t + \varepsilon_{it} \quad (6)$$

### 3.6 Theoretical Perspectives on ESG's Influence on Corporate Financial Performance

In a preliminary step, before highlighting theories of how ESG may impact investor returns, it is crucial to explore the diverse theories regarding the influence of ESG on a firm's financial performance. This conceptual framework aids in contextualising and understanding various viewpoints. It is pivotal in facilitating a more informed and structured analysis when subsequently examining how ESG factors may influence investor returns. Notably, two prominent and contrasting theories emerge: the shareholder and stakeholder models, complemented by additional relevant models.

#### 3.6.1 Negative ESG - Corporate Financial Performance Models

The trade-off hypothesis, dating back to Levitt (1958) and influenced by Milton Friedman's shareholder model, posits that a company's primary social responsibility is profit maximisation for shareholders. This perspective sees sustainability investments with scepticism, viewing them as depleting resources and potentially compromising financial performance. Friedman argued that allocating funds to social causes diverts resources from a company's core expertise, leading to inefficient wealth use and negative societal impacts (Friedman, 1970). According to the trade-off theory, companies engaging in ESG investments are expected to perform less favourably than other companies. This is attributed to the belief that such investments drain resources, incurring social costs perceived to outweigh potential gains.

Additionally, the managerial opportunism hypothesis posits that managers, incentivised by compensation schemes linked to short-term profits and stock prices, may prioritise personal objectives at the expense of both shareholders and stakeholders. Pursuing personal managerial goals could result in a negative relationship between financial and social performance, as strong financial performance may prompt managers to minimise social expenditures for short-term private gains. Conversely, during periods of weakened financial performance, managers might seek to offset and justify disappointing results by engaging in conspicuous social programs (Preston & O'Bannon, 1997).

---

<sup>4</sup> The operating profitability ratio used is the annual revenues minus cost of goods sold, interest expense, and selling, general, and administrative expense divided by the sum of book equity and minority interest.

<sup>5</sup> The investment ratio used is the change in total assets divided by previous year's total assets

### 3.6.2 Positive ESG - Corporate Financial Performance Models

Stakeholder theory suggests that ESG is positively associated with financial performance. Pioneered by scholars like Ed Freeman in 1984, this perspective argues that considering all stakeholders can lead to superior performance compared to exclusively prioritising shareholders (Freeman, 2010). In contrast to the idea that shareholder wealth maximisation should be the sole business objective, the instrumental stakeholder theory suggests that meeting key stakeholders' demands, including enhancing Corporate Social Responsibility (CSR), contributes to superior Corporate Financial Performance (CFP). Effective stakeholder management, along with meeting expectations, can provide firms with competitive advantages, such as a positive reputation and enduring relationships (Endrikat et al., 2014). In alignment with the stakeholder theory, companies engaged in ESG investments are expected to outperform their counterparts, as integrating environmental, social, and governance considerations is seen as fostering enhanced stakeholder relationships and enduring sustainable success.

Furthermore, two other prominent theories warrant attention. Firstly, Hart (1995) introduced the Natural Resource-Based View (NRBV), positing that a firm's adeptness in addressing challenges from the natural environment fosters the cultivation of distinctive and challenging-to-replicate organisational resources and capabilities. The NRBV proposes that a company can attain a competitive advantage and realise superior financial performance by efficiently managing and leveraging its natural resources in response to environmental challenges. This perspective underscores the significance of environmental sustainability and resource management in contributing to a firm's overall strategic advantage. Secondly, the Porter hypothesis posits that regulations promoting CSR can stimulate innovations, ultimately improving a firm's efficiency and competitiveness (Esty & Porter, 1998). Despite the time required for innovation development, firms are likely to collect continued benefits, as they stand a higher chance of securing a competitive advantage, thereby leading to improved financial performance.

Alternatively, some scholars propose a positive correlation, but with the causal link operating from financial performance to ESG performance, indicating a reverse direction in the relationship. This viewpoint, called the slack resource hypothesis or available funding theory, posits that firms embrace sustainable practices contingent on the availability of financial resources, suggesting that sustainability depends on financial backing (Preston & O'Bannon, 1997).

Thus, a synergistic relationship could exist, forming a virtuous circle. According to the stakeholder theory, high ESG performance may enhance financial performance. Subsequently, according to the theory of slack resources, better financial performance can lead to improved ESG performance, creating a positive bidirectional relationship. Conversely, a vicious cycle may occur, where a deficiency in ESG performance results in poor financial performance, leading to reduced ESG investments (Waddock & Graves, 1997).

### 3.7 Theoretical Framework for Assessing the Impact of ESG on Investor Returns

The preceding section presented arguments regarding how ESG factors may influence a company's financial performance, and this section highlights theories concerning how investors can or cannot use this ESG information to attain abnormal returns.

According to established theories, such as the Efficient Market Hypothesis, leveraging ESG information is expected to not lead to abnormal returns for investors. Given that ESG ratings are publicly available metrics, if markets are at least semi-efficient, all information embedded in these scores should already be incorporated into the market prices of companies. Therefore, if the prevailing consensus suggests that enhanced sustainability performance could positively impact future operational performance, resulting in increased cash flows and higher future valuation, or a reduction in operational risk, the asset's value is anticipated to rise. Consequently, an increase in ESG performance would lead to increased demand for the stock, leading to an immediate price adjustment rather than a future price increase. In this scenario, stock prices consistently reflect their fair value, and utilising public ESG information is not expected to result in improved or reduced financial performance for the investor.

Moreover, in line with the principles of Modern Portfolio Theory, implementing various ESG screens, whether positive or negative, may limit the pool of assets available for portfolio construction. This restriction could potentially result in lower risk-adjusted returns for sustainable portfolios, as they might not fully capitalise on the advantages of diversification. However, some scholars, such as Barnett & Salomon (2006), contend that, in contrast to the principles of MPT, a decrease in diversification opportunities does not inherently result in diminished risk-adjusted performance. They argue that the advantages of limiting the investment universe may outweigh the associated costs, given that the remaining assets are better managed and possess lower inherent risk. In essence, the companies that remain available for investors to choose from are inherently superior to the broader market.

Additionally, specific situations may give rise to market anomalies, and ESG could potentially be considered an anomaly if investors exhibit irrational behaviour. Pedersen et al. (2021) elucidated in their publication how high-ESG stocks can either outperform, underperform, or perform equivalently to low-ESG stocks. The authors initiate their analysis by categorising all investors into three groups: "ESG-unaware," "ESG-aware," and "ESG-motivated".

"ESG-aware" investors are rational actors seeking to optimise their risk-adjusted returns while acknowledging the potential impact of ESG factors on these companies. In a market with a significant presence of ESG-aware investors, the prices of high-ESG stocks are elevated to a level that accurately mirrors their fair value, akin to the efficiency expected in an optimal market.

"ESG-unaware" investors aim to optimise their risk-adjusted returns but lack awareness of ESG considerations. In situations with a substantial presence of ESG-unaware investors in the market, the prices of ESG stocks might not experience significant upward pressure, even if these stocks are more lucrative (assuming higher ESG scores lead to superior financial performance). Consequently, in markets dominated by ESG-unaware investors, high-ESG stocks may exhibit outperformance compared to low-ESG stocks.

"ESG-motivated" investors are individuals who are aware of ESG factors and systematically prefer investing in high-ESG stocks due to the personal utility they derive, even if it results in lower risk-adjusted returns. In instances where a significant number of ESG-motivated investors are present, the prices of high-ESG stocks may be bid up beyond their fair value, as these investors are willing to accept lower returns for a portfolio with robust ESG characteristics. Consequently, in such situations, high-ESG stocks may underperform due to inflated prices influenced by the preferences of ESG-motivated investors, whose behaviour is driven by behavioural considerations.

Hvidkjær (2017) introduces an additional rationale for stocks' over- or underperformance, attributing it to the stock market's underreaction to ESG information. The central argument centres on the stock market's inclination to undervalue intangible factors due to their complex nature and exclusion from a company's financial statements. Since ESG initiatives often fall into the category of intangibles, the stock market is prone to displaying an inadequate response to ESG information.

## **4 Literature Review of Empirical Studies**

### **4.1 Empirical Insights on the Relationship between ESG and Corporate Financial Performance**

The upcoming sections explore the existing empirical literature on the relationship between specific ESG pillars - environmental, social, and governance performance - and corporate financial performance. The aim is to understand how ESG factors contribute to a company's financial success. Moreover, the causal link is examined, exploring whether one type of performance influences another or if they demonstrate synergistic interactions. This section exclusively provides evidence on the efficacy of ESG decisions as sound business practices, refraining from making claims about whether investors can attain superior returns by investing in these firms. The examination of investor returns is addressed in Chapter 4.2, which specifically focuses on investor studies.

#### **4.1.1 Corporate Social Performance and Corporate Financial Performance**

The initial focus is on the social pillar, which boasts deep historical roots dating back to ethical investing in the 1700s. In a modern context, the social pillar rose to prominence during the 1960s, a period characterised by significant political movements such as civil rights, gender equality, and an increased emphasis on social responsibility and accountability, spurred partly by the anti-Vietnam War movement (Schueth, 2003). Over time, this pillar has evolved to encompass a broader spectrum of considerations.

Extensive research has been conducted on this pillar, including well-known studies such as the one by Orlitzky et al. (2003), which compiled various findings in an extensive meta-analysis. They argue that companies with high Corporate Social Performance (CSP) can enhance relationships with bankers and investors, potentially facilitating access to capital, attracting top-tier employees, and fostering goodwill among existing staff, thus contributing to improved Corporate Financial Performance (CFP). This study challenged the notion that CSP inherently contradicts shareholder wealth maximisation, as suggested by the shareholder theory. Examining more than three decades of empirical data, they illustrate a positive and bidirectional correlation between CSP and CFP across a broad spectrum of industries.

Allouche & Laroche (2005) extended Orlitzky et al.'s (2003) work by utilising a larger sample size, considering both U.S. and U.K. studies, examining social performance in various roles, and employing advanced meta-analytic methods to address selection and publication bias. Their findings also supported a positive relationship, with a stronger context observed in the U.K., indicating that this phenomenon is not limited to the U.S. Similar to Orlitzky et al. (2003), they observed that CSP reputation indices had a more pronounced effect on CFP than social disclosures. Social disclosures constitute an objective variable, reflecting how a firm has contributed value to stakeholders in the past,

providing a retrospective perspective. Conversely, reputation serves as a perceptual variable and constitutes a predictive element, representing the anticipation of how the firm is expected to deliver value in the future (de Quevedo-Puente et al., 2007).

Similarly, Clark & Viehs (2014) examined various studies on CSP and concluded that the social dimension of the ESG universe generally exerts a positive influence on CFP. However, they highlighted a significant gap in the literature, noting the lack of examination of critical corporate social responsibility behaviours such as respect for human rights, socially responsible advertising campaigns, and worker safety.

Meta-analytical studies confirm a bidirectional relationship between CSP and CFP, though the relationship may not be strictly linear. Barnett & Salomon (2012) found a U-shaped pattern, where the highest CSP leads to the highest CFP. They suggest that socially responsible practices incur costs, but the positive impact on stakeholder relations, offsetting these costs, requires substantial Stakeholder Influence Capital (SIC). In instances of weak social performance, anticipated benefits may not materialise, resulting in a negative CSP-CFP relationship. Increased investment in social issues without improving stakeholder relations leads to more losses. As firms accumulate SIC through heightened social performance, an inflexion point occurs, enhancing the ability to convert social investments into positive financial returns. Companies with the highest social performance demonstrate superior capabilities, contributing to an upward trajectory in financial performance.

#### **4.1.2 Corporate Environmental Performance and Corporate Financial Performance**

A growing body of research has pivoted its attention from the social dimension to the environmental realm, particularly focusing on Corporate Environmental Performance (CEP). This shift reflects the mounting concerns surrounding climate change and environmental degradation. These studies carry notable importance due to their technical nature, demanding specific firm capabilities, significant capital investments, compliance with regulatory frameworks, and adherence to distinct reporting criteria.

Endrikat et al. (2014) undertook an extensive meta-analysis of 149 individual studies, considering both the causality direction and the multidimensionality of focal constructs. They distinguished between process-based CEP, which signifies a firm's internal initiatives to tackle environmental issues, and outcome-based CEP, which represents these efforts' observable and quantifiable results. Logically, process-based CEP should lead to outcome-based CEP. The study revealed a positive association, rejecting the argument of a trade-off between CEP and CFP. This positive relationship was found to be bidirectional, with CEP positively influencing CFP and a partially positive link observed in the reverse direction.

When distinguishing between process and outcome-based CEP, it was observed that process-based CEP is more likely to correlate with accounting-based CFP, whereas

outcome-based CEP is more likely to align with market-based CFP and accounting-based CFP. This suggests that the market tends to reward a firm's environmental performance when it is associated with tangible and measurable outcomes, as opposed to environmental processes, which might be perceived as less reliable measures or potential greenwashing attempts. Moreover, the study found that a proactive approach to addressing environmental concerns within a firm is more likely to be linked with enhanced financial performance than reactive approaches. This observation aligns with the Natural Resource-Based View (NRBV), emphasising the significance of a company's proactive management of natural resources for achieving a competitive advantage and superior financial outcomes (Hart, 1995).

Hang et al. (2019) also investigated the bidirectional relationship between CEP and CFP through a meta-analysis of 142 primary empirical studies. The study underscored the causal effect, emphasising its temporal dependency. Over a short duration (less than one year), it validated the slack resource hypothesis, implying that financial resources can enhance a firm's environmental performance. On the contrary, a high level of CEP does not yield an immediate impact on CFP. However, the firm realises substantial long-term benefits, consistent with the principles of the Porter hypothesis. Consequently, a company needs resources to support initial investments for improving environmental performance, with the returns materialising in the long term. Furthermore, consistent with the findings of Endrikat et al. (2014), they observed that reactive environmental investments yield smaller financial performance gains than proactive ones.

When scrutinising the relationship dynamics, Trumpp & Guenther (2017) identified a U-shaped relationship between CEP and CFP, resembling the pattern observed in social performance by Barnett & Salomon (2012). They propose the "too-little-of-a-good-thing" framework, suggesting that being environmentally friendly pays off only after surpassing a certain minimum level of CEP. When using linear relationships, they find either a significantly positive result or insignificant results, emphasising that linear models struggle to capture the complex relationship between CEP and CFP influenced by various hypotheses like trade-off, managerial opportunism, stakeholder, NRBV, etc., potentially leading to misleading results.

#### **4.1.3 Corporate Governance and Corporate Financial Performance**

Governance and social responsibility are closely related in the business context. Governance pertains to the framework of regulations, customs, and procedures through which a company is guided and overseen. On the other hand, social responsibility involves a company's commitment to acting ethically and contributing positively to society and the environment (Jamali et al., 2008). Consequently, Corporate Governance (CG) has frequently been examined as an integral aspect of Corporate Social Responsibility, as highlighted by Sahut et al. (2019). Despite this, certain studies have explored the direct relationship between CG and CFP.

Findings related to the relationship between CG and CFP can be mostly explained through the insights provided by agency theory (Jensen & Meckling, 1976). This theory posits a potential conflict of interest between principals (shareholders) and agents (management). Effective governance mechanisms, such as robust boards and transparent policies, are crucial in aligning the interests of management with those of shareholders. Accountability and monitoring of executives and directors reduce the likelihood of agency problems, where management pursues goals diverging from shareholder interests.

Research exploring this relationship includes Bhagat & Bolton (2008), revealing a significant positive correlation between enhanced governance, measured through different indices, board members' stock ownership, CEO-chair separation, and both concurrent and subsequent operational performance. However, none of these governance measures correlate with future stock market performance. In a later study, Bhagat & Bolton (2019) extended the sample period from 2002 to 2016. Their findings indicate that director stock ownership consistently and positively correlates with future corporate performance. Notably, within the financial sector, particularly in the sub-sample of banks, director stock ownership is positively associated with future bank performance and negatively correlated with future bank risk.

In their analysis, Paniagua et al. (2018) explored the relationship by examining factors such as the number of board members, ownership dispersion, and the cost of ownership. While the OLS regression failed to provide conclusive results, the Poisson regression demonstrated a negative association between an increase in the number of board members, ownership dispersion, or the cost of ownership (dividends) and a firm's financial performance. The increase in board members and ownership dispersion might complicate the alignment of interests between management and shareholders, thus leading to worse corporate governance, which may explain the reduction in financial performance.

Consistent with previously mentioned studies, Alley et al. (2016) demonstrated a positive relationship. Notably, board expertise, composition, and managerial skills contributed to improved financial performance indicators, such as return on equity, return on assets, and net profit margin. On the flip side, the size of the board and audit committee did not show a positive impact. They could even undermine financial performance, as supported by the findings of Paniagua et al. (2018). It's essential to highlight that this relationship is unidirectional, implying that corporate governance positively influences CFP, but the reverse causation is not observed.

Rossi et al. (2015) uncovered a positive relationship with return on equity but a negative association with firm market value. The authors argue that effective governance cultivates an efficient and non-speculative management approach, deterring actions that could detrimentally impact shareholder wealth. While the firm's value may be lower in the short term, this strategy maximises shareholder value in the long term by avoiding speculative operations by examining management compensation schemes and safeguarding shareholder and stakeholder rights.

#### **4.1.4 Summary: ESG and Corporate Financial Performance**

In essence, these studies collectively affirm that each sustainability pillar - Environmental, Social, and Governance - positively influences the performance of individual companies. This collective impact underscores the multifaceted nature of sustainability and its potential to enhance various aspects of CFP.

Considering all the pillars collectively, it is unsurprising that meta-studies consistently reveal a robust positive relationship between ESG factors and company financial performance. Among the most cited works is the study by Friede et al. (2015), which synthesises findings from over 2,200 diverse studies. In more than half of the cases (56.7%), a positive relationship is identified, with only 9.2% showcasing a negative relationship. This trend aligns with other recent meta-studies, including Atz et al. (2023), reporting a positive relationship in 60% of cases, and Whelan et al. (2021) and Sinha et al. (2019), indicating a positive relationship in 58% and 66% of the studies, respectively. Friede et al. (2015) attempted to identify the specific pillars leading to this higher performance. They discovered that the most positive relationships occur when companies exhibit strong governance, followed by environmental and social performance. Nevertheless, no substantial differences were observed, as each distinct pillar demonstrated a positive relationship in over 50% of cases.

The extensive analyses conducted offer compelling evidence supporting a positive correlation between ESG factors and financial performance at the company level. Furthermore, this relationship may manifest bidirectionally under various circumstances. Each pillar holds its unique significance. Regarding environmental investments, the key is taking proactive environmental measures, as financial performance is most positively influenced by proactive rather than reactive approaches. In the societal realm, the paramount importance lies in cultivating a strong reputation. As for governance, the critical focus is on aligning the interests of shareholders and management. These relationships are influenced by numerous variables, suggesting that the dynamics may not conform to a linear pattern. In several instances, the relationships exhibit a U-shaped trajectory, implying that the benefits of a robust ESG profile might become more pronounced beyond a certain threshold. However, it is important to emphasise that this does not ensure consistent abnormal returns from investing in these companies. Many of these studies rely on accounting measures, and any performance improvement may already be factored into market prices.

## **4.2 Empirical Investor Studies on the Relationship between ESG and Financial Performance**

In contrast to individual company studies, investor studies focus on assessing the impact of ESG on financial performance not at the level of specific firm data but rather through the examination of aggregated firm performance within virtual portfolios and financial products like mutual funds or indices. These studies, often termed portfolio studies, are particularly crucial for investors seeking to integrate ESG information into

their portfolio allocation strategies. Their primary objective is to determine whether it is feasible to achieve abnormal returns when investing based on ESG criteria. Specifically, these studies investigate the potential to outperform the market and generate "alpha" by leveraging ESG information. This chapter is grounded on the theories explored in Chapter 3.7.

#### **4.2.1 ESG Screening at Portfolio Level**

ESG screening can come with additional costs compared to conventional investing, as outlined by Becchetti et al. (2015). These costs encompass the expenses associated with obtaining specific CSR information about stocks available for investment, the costs linked to forgoing diversification opportunities due to imposed investment constraints according to the Modern Portfolio Theory, and timing costs. The latter comes into play when a fund manager is obligated by the fund's rules to divest a stock if the company undergoes a behavioural change, losing its social responsibility characteristics. This requirement might compel the fund manager to sell the stock, even if the company is anticipated to have promising return prospects, resulting in a suboptimal transaction and leading to heightened turnover costs. For instance, Wimmer (2013) analysed the persistence of ESG scores in socially responsible mutual funds and found that ESG scores persist for approximately two years but fade away after three years.

In principle, the construction of ESG-screened portfolios is designed to reduce overall ESG risk within the portfolio by excluding low-ESG score components from the pool of eligible selections. When investors effectively implement this screening process, they anticipate that ESG-screened portfolios will be safeguarded against losses caused by ESG-related events and have the potential to generate higher realised alphas than portfolios that are not screened. This aligns with Barnett and Salomon's (2006) argument that the benefits of screening may outweigh its costs.

Jin (2022) conducted a theoretical examination in his study, suggesting that according to Modern Portfolio Theory, investors can lower the specific risk of portfolios to extremely low levels through diversification. A well-diversified portfolio is expected to deliver returns that align with their systematic risk. Consequently, this implies that ESG screening at the portfolio level may become redundant if ESG-related risks at the individual constituent level can be fully mitigated through diversification. Furthermore, ESG screening can raise costs and specific risks by limiting diversification opportunities. This theoretical analysis suggests that engaging in screening may not yield benefits, as ESG risk is already diversified and may incur additional costs.

Several studies have examined how ESG constraints affect the efficient frontier and have attempted to quantify the corresponding costs. Geczy et al. (2005) discovered that Socially Responsible Investment (SRI) constraints incurred a marginal cost of approximately 5 to 10 basis points per month, based on an analysis of U.S. equity funds spanning from 1969 to 2001. Herzel et al. (2012) examined S&P500 components from 1993 to 2008 and observed that SRI screening had a negligible impact on Sharpe ratios. Similarly, Qi & Li

(2020) explored the effects of these constraints on the Dow Jones Industrial Average Index from 2004 to 2013, revealing no substantial difference in portfolio performance when ESG constraints were applied.

Pedersen et al. (2021) took a unique approach by investigating the impact of various screening strategies on the S&P500 and the corresponding maximum achievable Sharpe Ratio. Instead of illustrating a conventional mean-variance frontier, they depicted an ESG-Sharpe Ratio frontier. The findings suggested that the mean-variance efficient portfolio has a specific ESG ratio. Attempting to raise the ESG ratio resulted in a decline in the Sharpe ratio, which became more pronounced when a higher ESG score was required.

Consequently, while there is an associated cost with ESG constraints in investment portfolios, empirical findings demonstrate some variability, and these costs are generally not excessively high. As the pool of available stocks not subject to these constraints has grown, the potential for diversification has likely increased, thereby reducing these associated costs.

Additional research has looked into screening intensity, proposing that merely considering whether screening is employed or not is insufficient for investigating the impact. Instead, it proposes focusing on the number of screens employed in the screening process. For example, Lee et al. (2010) identified a curvilinear relationship between screening intensity and risk. With up to 6 screens applied, the overall risk decreased due to lower systematic risk. However, increasing the number of screens beyond this point led to increased risk, as fund managers encountered challenges in identifying lower beta stocks. Intriguingly, idiosyncratic risk remained unaffected, indicating there were still enough stocks to diversify firm-specific risks. While screening reduced overall risk, it adversely impacted risk-adjusted performance with each additional screen applied, suggesting that a balanced approach to screening is advisable.

Jin (2022) offered a unique perspective by examining the concentration level of ESG screening rather than its intensity. Their analysis excluded low ESG-score funds from a portfolio comprising U.S. Equity Funds. The findings revealed that portfolios with a higher concentration of ESG considerations displayed less risky performance, primarily due to reduced systematic risk. Additionally, these portfolios exhibited enhanced risk and return performance, influenced by diverse style factor exposures. Interestingly, they observed a marginal decrease in specific risks up to a concentration level where 50% of low ESG funds were screened out, but beyond this threshold, specific risks increased rapidly. Finally, portfolios exhibiting higher concentration levels tended to undergo increased turnover rates. He emphasised that screening can profoundly influence investment performance, even at the portfolio level, after mitigating individual ESG risks through diversification. Thus, the concentration level of ESG screening should be used as a search parameter to strike a balance between costs and benefits.

These articles underscore that ESG screening can substantially influence investment performance, even after individual ESG-related risks within constituent investments

have been reduced through diversification. This reaffirms the rationale for implementing ESG considerations at the portfolio level. Nevertheless, it also emphasises the importance of thoughtful decision-making in implementing ESG criteria at the portfolio level. It's essential to recognise that applying these ESG screens may either enhance or reduce the performance of socially responsible funds, depending on the intensity or the type of screens employed (Renneboog et al., 2008). According to Barnett and Salomon (2006), the connection between performance and ESG depends on how fund managers employ ESG criteria. They assert that positive returns are associated with using ESG criteria to adjust the weight of portfolios, particularly steering away from poor ESG companies rather than outright exclusion. A prevailing observation is that a positive correlation with financial performance is possible only when fund managers have the flexibility to deviate from rigid ESG screening criteria.

#### **4.2.2 Empirical Findings of Portfolio Studies**

Several studies have broadened their focus from individual company analyses to portfolio studies. This expansion is justified by the prevailing evidence supporting the positive impact of ESG on company financial performance and the recognised potential for ESG integration to generate favourable outcomes at the portfolio level, provided it is executed with due diligence. Subsequently, the following studies assessed the relationship between ESG factors and financial performance on a portfolio level, with financial performance gauged through various return measures.

In the initial set of empirical findings, meta-studies are examined. Friede et al.'s (2015) meta-analysis, which primarily highlighted a positive relationship between ESG and financial performance at the company level, revealed a considerably weakened relationship when analysed at the portfolio level. Out of 155 portfolio-focused studies, only 15.5% displayed a positive association, while 11% indicated a negative one, leaving the rest with inconclusive or neutral results. Azt et al. (2023) arrived at a similar conclusion, noting a less distinct positive relationship at the portfolio level, with 38% showing a positive correlation and 13% a negative one. Likewise, Whelan et al. (2021) identified a positive relationship in 33% of cases and a negative one in 13%. Shifting the focus from company-level to portfolio-level studies, they observed a decline in the proportion of positive relationships and an increase in the percentage of neutral, mixed, and negative findings. Additionally, Azt et al. (2023) discovered that both actively and passively managed ESG funds are likely to have a similar positive impact on financial performance. Furthermore, they emphasised the critical role of the portfolio management strategy. While simple negative screening and divesting showed neutral or mixed findings, ESG integration exhibited a positive relationship in 59% of the studies. Similar conclusions were drawn by Whelan et al. (2021), highlighting the limitations of relying solely on rigid ESG screening. This underscores the importance of avoiding a simplistic approach of excluding certain industries or stocks and emphasises the need to explore alternative methods for ESG integration.

#### ***4.2.2.1 Empirical Evidence of Equal Performance***

In the initial studies exemplified by Bello (2005), the focus was primarily on the performance of socially responsible mutual funds compared to conventional mutual funds. Adopting a matched-pair methodology, Bello paired two conventional mutual funds with each socially responsible mutual fund, considering similar net assets. Financial performance was assessed between 1994 and 2001 using metrics such as Jensen's alpha, Sharpe ratio, and excess standard deviation-adjusted returns. Despite the expectation that the screened portfolio would underperform compared to the conventional one, both portfolios demonstrated equal risk-adjusted returns.

Other studies, such as Humphrey et al. (2012), took a different approach by concentrating on individual stocks instead of mutual funds. They constructed portfolios comprising high and low-scoring companies based on ESG criteria of a U.K. sample from 2002 to 2010. ESG scores were provided by Sustainability Asset Management Group GmbH. Employing a Carhart 4-factor model, the study's findings indicated no significant difference in risk-adjusted returns between low- and high-ESG score portfolios, concluding that the U.K. markets are efficient. Consequently, the study suggested that U.K. investors could incorporate ESG considerations without fearing lower returns, albeit without expecting higher returns.

Halbritter & Dorfleitner (2015) focused on the U.S. stock market spanning from 1991 to 2012. They established top and bottom decile portfolios, equal and market-weighted, using three different rating providers to overcome challenges with diverging ESG scores (Berg et al., 2022). Employing the Carhart 4-factor model and cross-sectional Fama & Macbeth (1973) regressions, they assessed the performance of these ESG portfolios. Their results indicated no significant return difference between companies with high and low ESG ratings for each rating provider. While Fama & Macbeth (1973) regressions uncovered some influence of various ESG variables, capitalising on this relationship proved challenging for investors. The magnitude and direction of the impact significantly hinged on the rating provider, the company sample, and the specific subperiod.

Milonas et al. (2022) utilised a matched pair methodology to construct portfolios comprising high and low-scoring ESG funds. Their sample included 80 U.S. and 64 European mutual funds observed between 2017 and 2021. Performance evaluation incorporated metrics such as alpha, Sharpe ratio, Treynor ratio, excess daily returns, and the Fama and French 5-factor model. The study revealed no significant difference in risk-adjusted returns between ESG and non-ESG funds, even though ESG funds demonstrated slightly higher returns.

Teti et al. (2023) investigated the relationship by forming portfolios with high and low Refinitiv ESG combined scores using the STOXX Europe 600 constituents from 2016 to 2021. The study revealed compelling evidence indicating that the portfolio in the bottom decile yielded negative alpha. Conversely, a long-short strategy, involving purchasing the high ESG portfolio and selling the low ESG portfolio, resulted in some positive abnormal

returns across three prominent asset pricing models, although not statistically significant.

De Spiegeleer et al. (2023) examined the consequences of integrating ESG criteria into the allocation of equity portfolios, focusing on assets in the STOXX Europe 600 and the Russell 1000 index. The analysis did not reveal conclusive evidence supporting improved performance for portfolios with either high or low ESG scores. Specifically, when assessing risk-adjusted returns in both countries from 2010 to 2020, portfolios with high ESG ratings demonstrated underperformance. However, this underperformance was predominantly attributed to the excessive risk-adjusted returns of low ESG-rated portfolios in the initial years of the observation period. In contrast, when considering a more recent observation period, commencing from 2016/2017 onwards, portfolios with high ESG ratings exhibited notably superior risk-adjusted returns. Additionally, the outcomes were highly contingent on the ESG rating agency data employed, with the primary findings being expressed in MSCI ESG ratings. Consequently, the researchers concluded that the influence of ESG rating constraints on a portfolio is subject to variations based on the choice of the rating agency, the specific market universe, the precise investment period, and other relevant factors.

#### ***4.2.2.2 Empirical Evidence of Outperformance***

Early studies conducted by Kempf & Osthoff (2007) and Statman & Glushkov (2009) analysed the impact of social responsibility on portfolio performance using MSCI KLD ratings for constituents from the S&P500 and DS400. Both studies concluded that significant abnormal returns could be achieved, particularly when employing a best-in-class approach, while negative or positive screening did not yield similar results. Kempf & Osthoff (2007) further demonstrated the persistence of abnormal returns even after considering additional transaction costs. Statman & Glushkov (2009) emphasised the drawbacks of excluding stocks from rejected companies, suggesting that the best-in-class method provided a strategic approach to constructing portfolios based on companies with high social responsibility scores on characteristics like community, employee relations, and the environment without automatically excluding those associated with sin industries.

Steen et al. (2020) undertook an investigation into the relationship between Morningstar's Sustainability Ratings and the performance of 146 mutual funds domiciliated in Norway between 2014 and 2018. Utilising the Fama and French 3-factor model, they analysed various ESG quintiles and found no abnormal risk-adjusted returns in the overall dataset. Nevertheless, they conducted a separate analysis focused on European-categorised funds to address geographical bias. In this refined analysis, they identified significantly higher returns and positive alphas within the top ESG quintile, a distinction attributed to these funds' superior social and governance scores.

Abate et al. (2021) examined the performance of 634 European mutual funds over the period from 2014 to 2019. They uncovered that funds in the top decile of Morningstar

Sustainability Rating consistently demonstrated significantly higher average Data-Envelopment-Analysis (DEA) efficiency compared to their counterparts in the bottom decile. Interestingly, despite engaging in more intricate screening activities and incurring associated operating costs, the highly-rated funds managed to maintain lower ongoing charges. This aligns with similar observations highlighted by ESMA (2022b).

Amon et al. (2021) adopted a more realistic approach in constructing ESG portfolios to assess their risk-adjusted returns relative to the benchmark. Instead of simply categorising portfolios into high and low ESG percentiles, they devised ESG-weighted portfolios that adjust asset allocation toward assets demonstrating strong ESG performance. These portfolios were constructed based on the constituents of the S&P 500 and STOXX Europe 600 indices as of 2018. Employing the Fama & French 5-factor model, the study revealed that an ESG-weighted portfolio outperformed a value-weighted portfolio while also exhibiting superior ESG performance between 2005 and 2018. In comparison to an equally weighted portfolio, the ESG-weighted portfolio delivered comparable risk-adjusted returns but excelled in terms of ESG performance.

#### ***4.2.2.3 Empirical Evidence of Underperformance***

In the existing body of literature, the instances of studies indicating underperformance are notably fewer than those indicating equal or outperformance. This trend is evident in the meta-analysis carried out by Friede et al. (2015), Whelan et al. (2021) and Azt et al. (2023).

Bauer et al. (2005) conducted an analysis of ethical mutual funds from the U.K., the U.S., and Germany, spanning the period from 1990 to 2001. They employed the Carhart 4-factor model to compare the returns of ethical funds with those of conventional funds. The overall results did not reveal any significant difference in financial performance between ethical and conventional funds. The authors further examined whether the financial performance of the funds varied over time by dividing the sample into three chronological periods. The findings indicated that ethical funds significantly underperformed at the beginning of the period, while towards the end, they performed on par with conventional funds. The study suggested that ethical funds might have undergone a learning curve over time, with older funds catching up after a period of pronounced underperformance while younger funds still exhibited underperformance.

Das et al. (2018) investigated the risk-adjusted performance of socially responsible mutual funds domiciled in the U.S. with high and low Morningstar Sustainability Ratings from 2005 to 2016. Their methodology involved a multi-factor regression, with the dependent variable being the Sharpe ratio and controlling for factors such as management tenure, expense ratio, age, and fund size. They identified a negative relationship, with an exception noted during the Global Financial Crisis, indicating that ESG primarily functions as an insurance mechanism.

Bannier et al. (2019) constructed portfolios with high and low Refinitiv ESG scores using available data for U.S. and European companies spanning from 2003 to 2017. Employing a Carhart 4-factor model, they determined that the high ESG portfolio did not generate abnormal risk-adjusted returns compared to the market benchmark. However, implementing a long-short strategy resulted in abnormal negative returns, indicating that firms with high ESG scores did not yield negative returns, while those with low ESG scores offered a highly significant excess return.

### **4.2.3 Conclusion & Reasons for Divergence**

As indicated by the meta-analysis, the results are varied, with the majority suggesting no statistically significant relationship, aligning with the hypothesis of overall market efficiency. Certain evidence suggests that basic negative screening activities may lead to underperformance compared to the market (Bauer et al., 2005; Kempf & Osthoff, 2007; Statman & Glushkov, 2009; Lee et al., 2010). Moreover, studies have shown that top ESG companies may not necessarily outperform, while low ESG companies could either underperform (Teti et al., 2023) or overperform (Bannier et al., 2019), creating opportunities for abnormal returns through a long-short strategy. It's essential to consider variations in rating providers, including Refinitiv, Morningstar, and MSCI, diverse timeframes spanning from 1990 to 2021, different geographical regions (mainly the United States and Europe), and various methodologies such as active and passive managed funds, custom portfolios, and a range of statistical approaches like multi-factor regression, Data Envelopment Analysis, Sharpe ratio, among others. These are critics as well addressed by numerous researchers, notably Atz et al. (2023), who suggest various reasons to account for the mixed findings in portfolio studies, among which three appear particularly significant:

Firstly, practical limitations with data quality and consistency issues in ESG scores play a crucial role (Atz et al., 2023). With over 10 to 15 major ESG rating providers and numerous smaller ones, the choice of the rating agency used in a study can significantly impact the results (ESMA, 2022a). Given the consistency issues discussed in Chapter 2.2, researchers should exercise caution in selecting the rating agency to ensure the robustness of their findings.

Secondly, many researchers tend to combine various portfolio management strategies without distinguishing between them (Atz et al., 2023). This issue is particularly evident in mutual fund studies, where conducting a fair matched pair analysis becomes challenging. Simple regression analysis across different mutual funds combines varied strategies and allocations, creating complexity affecting accurate comparisons. For instance, ethical investment funds may not aim to outperform the market but are categorised as an ESG strategy.

Lastly, Atz et al. (2023) argue that ESG investing offers asymmetrical advantages, notably in terms of downside protection, making its benefits most apparent under specific circumstances. This asymmetry can lead to mixed findings in portfolio studies as the

effectiveness of ESG strategies may vary depending on market conditions, economic cycles, and the occurrence of specific events. During periods of market turbulence or economic downturns, the risk mitigation aspects of ESG investing may become more pronounced, contributing to positive results. However, in more stable or bullish market conditions, the impact of ESG factors on financial performance may be less discernible, leading to a diverse range of outcomes in different studies.

The relationship between ESG factors and financial performance remains a nuanced and multifaceted area of study, with findings varying across different methodologies, periods, and geographic regions. The impact of ESG factors on financial performance is a complex and evolving field, offering ample opportunities for further research and exploration. Overall, the results are mixed, and drawing definitive conclusions is challenging, underscoring the significance of contextual considerations.

#### **4.2.4 Insurance Effect during Economic Downturns**

In addition to exploring the broader connection between ESG ratings and financial performance, numerous studies have explored whether ESG can function as a safeguard against firm-specific risks, as ESG might offer asymmetrical benefits in certain conditions. Over the past two decades, three significant stock market crashes disrupted the world: the Dot-com Bubble burst from 2000 to 2002, the Great Financial Crisis (GFC) of 2007-2008, and the more recent market crash in 2020 linked to the COVID-19 pandemic. Two overarching hypotheses could explain how the COVID-19 pandemic and other crises might have influenced financial performance.

Firstly, Albuquerque et al. (2019) proposed that companies invest in ESG practices to distinguish their products, thereby cultivating a more loyal customer base with reduced sensitivity to price changes. This increased customer loyalty enables ESG-oriented firms to maintain higher price levels and profit margins. As the COVID-19 shock disrupted customer demand, it is hypothesised that firms with a strong commitment to ESG principles would benefit from this loyalty, potentially leading to more favourable stock performance.

Secondly, Bollen (2007) and Renneboog et al. (2011) suggested that investors who prioritise ESG considerations are less swayed by the performance of SRI funds compared to conventional mutual funds. The COVID-19 crisis influenced investors' risk attitudes, and ESG investors were perceived as more resilient in the face of market turmoil than those invested in conventional stocks. If the COVID-19 shock prompted a flight of investors from the market, albeit to a lesser extent among ESG investors, prices of ESG stocks would not experience as significant a decline in comparison to non-ESG stocks.

Exploring the role of ESG during significant financial crises, studies such as the one conducted by Lins et al. (2017) shed light on the performance of companies with strong CSR practices compared to those with lower CSR practices. This research revealed that, during both the Dot-com Bubble and the Global Financial Crisis, firms prioritising CSR

demonstrated superior performance. This trend persisted even after accounting for firm characteristics and risk factors. The study emphasises the idea that trust established between a company and its stakeholders and investors, fostered by investments in social capital, becomes particularly beneficial when overall trust in corporations and markets is diminished. The enhanced returns were attributed to higher profitability, margins, sales growth, and employee productivity among companies with stronger CSR practices. Importantly, this outperformance was specifically observed during these crises rather than in the subsequent recovery periods, suggesting that CSR might function as a form of insurance during market downturns. Investors may be willing to accept lower returns in normal times, anticipating the resilience it offers during crises (Engle et al., 2020). Numerous studies, including the work by Becchetti et al. (2015), reached comparable findings. While there was no discernible distinction in risk-adjusted performance between sustainable and non-sustainable investment funds from 1992 to 2012, there was evidence of enhanced financial performance for sustainable funds during the Global Financial Crisis, though not necessarily during the Dot-com Bubble. Furthermore, Bannier et al. (2019) highlighted the insurance-like function between 2003 and 2017. Their findings revealed that low-ESG companies outperformed high-ESG companies. However, high-ESG companies exhibited less downside risk and reinforced firm stability, contributing to the explanation of why low-ESG companies needed to offer investors higher risk premia. The mitigating impact of ESG activities on firm-specific risk was more pronounced during periods of higher market volatility, particularly during financial crises, but only in the U.S. market. No such correlation was observed in the European market. Similarly, ESG efforts significantly reduced a firm's market-based default risk in the U.S. but not in Europe.

The crisis triggered by the COVID-19 pandemic presented an unparalleled shock to financial markets, characterised by its exogenous origin tied to public health concerns rather than economic factors. This differentiates it from previous crises, such as the Global Financial Crisis and the Dot-com bubble. In examining the impact of COVID-19 on stock returns, Albuquerque et al. (2020) found that stocks with elevated environmental and social ratings exhibited markedly higher returns, lower volatility, and increased operating profit margins in the first quarter of 2020. The study proposes that both reduced investor flight and customer loyalty influenced the return performance of highly sustainable firms. However, the findings underscore the heightened significance of customer loyalty in explaining stock returns, particularly when firms boast both high ESG ratings and a substantial commitment to advertising.

In their analysis of the COVID-19 financial crisis in China, Broadstock et al. (2021) examined the role of ESG performance. Similar to findings by Albuquerque et al. (2020), Broadstock and his team observed that ESG-focused firms exhibited greater resilience in terms of stock price reactions to the COVID-19 crisis, thus leading to increased risk-adjusted performance during that time. However, when considering the overall data timeline, a negative relationship emerged, underscoring the insurance-like function of ESG. Furthermore, the study delved into the significance of individual E, S, and G scores. It concluded that higher E and G scores positively impacted performance, primarily due to their contributions to enhanced resiliency and financial stability when facing external

shocks like COVID-19. Conversely, higher S scores were associated with a negative impact on performance. This was attributed to their correlation with a heightened commitment to retaining or furloughing staff during crises - a reflection of a socially responsible approach that prioritises employee well-being over immediate cost-cutting measures such as layoffs.

Similarly, Petridis et al. (2023) conducted a comprehensive study involving a large dataset of 17,961 mutual funds. By integrating ESG data from Refinitiv and financial data from the latter part of the COVID-19 pandemic, they applied the DEA method, considering inputs related to expenses and risks and outputs related to returns for mutual funds. Their findings revealed that mutual funds with higher ESG controversy scores, indicative of fewer ESG controversies, exhibited superior financial performance. This trend persisted irrespective of the geographic regions in which these funds were invested.

In summary, historical evidence underscores that socially responsible investments tend to provide a form of downside protection during periods of market turbulence, effectively serving as a form of insurance. While the literature may present varying perspectives on the broader connection between sustainability and financial performance for an investor, there appears to be a consensus that a positive relationship tends to prevail during times of crisis. This perspective is supported by meta-analyses conducted by Whelan et al. (2021) and Azt et al. (2023). Furthermore, the significance of different sustainability factors can vary depending on the nature of the crisis, as exemplified by the recent COVID-19 crisis, where excessive emphasis on social score could harm investment performance.

## 5 Data & Methodology

### 5.1 Research Approach

In light of previous research indicating a predominantly positive relationship between ESG and financial performance at the firm level but nuanced findings at the portfolio level, this thesis examined this relationship within a portfolio context. The standard procedure for examining this relationship typically includes a quantitative analysis that constructs portfolios based on ESG ratings. Following this, back-testing is employed to compare the performance of these portfolios against benchmarks or other portfolios that do not integrate ESG information. This thesis employed a quantitative research methodology to address the central research question, leveraging statistical techniques to probe the connection between ESG considerations and financial performance. The initial step involved the construction of diverse portfolios, each characterised by unique ESG criteria. The subsequent step involved subjecting these portfolios to statistical analyses and multi-factor regressions. This study fundamentally questioned the Efficient Market Hypothesis by examining whether risk-adjusted returns are affected by ESG performance and whether ESG anomalies exist. The initial hypothesis, aligned with the efficient market theory, posits that no alphas should exist between different portfolios. However, this hypothesis is subject to confirmation or rejection based on the empirical evidence obtained from this research. Therefore, the central objective of this investigation was to extract insights into the relationship between ESG factors and financial performance. This involved leveraging well-established theories related to the maximisation of shareholder and stakeholder welfare, portfolio choice optimisation, and asset pricing models aimed at explaining stock price variations. Additionally, the study incorporated insights from prior scientific articles to enrich the understanding of a potential relationship.

In contrast to the returns-based approach commonly employed in studies comparing high and low-ESG-scoring mutual funds (Humphrey et al., 2012; Das et al., 2018; Steen et al., 2020; Milonas et al., 2022; Abate et al., 2021; Petridis et al., 2023), this study adopted a holdings-based approach (bottom-up approach). This method entails a thorough analysis of the ESG ratings of individual assets within an index or market, leading to the creation of custom portfolios. Examples include constructing portfolios comprising high and low-ESG-rated companies (Halbritter & Dorfleitner, 2015; Bannier et al., 2019; De Spiegeleer et al., 2023; Teti et al., 2023) or creating portfolios tilted towards high-ESG companies, such as ESG-weighted portfolios (Amon et al., 2021).

Several rationales support this decision. Firstly, Refinitiv provides current ESG ratings for mutual funds but lacks historical ESG data, which is currently only offered by Morningstar. The lack of access to Morningstar data presents challenges when analysing more extended reference periods due to potential changes in fund allocations and ESG ratings without the ability to rebalance the portfolios. Instead, the alternative involves creating custom portfolios based on companies' historical ESG data, which is provided by Refinitiv. Secondly, aggregated firm data allows a more transparent analysis of how the

financial performance of portfolios is influenced by their ESG characteristics. This approach avoids the added complexities of mutual fund dynamics, such as management fees, limited transparency regarding portfolio composition, differing time horizons, and distinct investment strategies. A bottom-up approach enables a direct response to the critique raised by Azt et al. (2023) concerning the challenge of comparing funds that employ different investment approaches.

Challenges related to utilising holding data are addressed by Liu & Strong (2008). They contend that portfolios constructed by researchers frequently comprise portfolios that would not be seriously considered ex-ante. Additionally, they underscore the impact of transaction costs on such portfolios, making them less practical as investment vehicles and potentially leading to misleading statistical inferences. To address these challenges, this paper consistently applied the holding approach to ensure a reliable comparison across all portfolios. Additionally, the rebalancing process, vital for accounting for annual ESG variations, occurred at limited intervals - specifically, once a year upon the release of new ESG data. This aligns with the common practice of quarterly or annual rebalancing for the weight of constituents in exchange-traded funds. Transaction costs were also factored in at a later stage. Considering that a typical investor would not construct a portfolio comprising solely 20% of the highest ESG-rated stocks, various portfolios were crafted, including naive ones like top and bottom quintile portfolios, alongside ESG-tilted portfolios inspired by the approach of Amon et al. (2021).

## 5.2 Data

### 5.2.1 Stock Data and Studied Time Period

Like most studies, this paper focused on constituents from the S&P 500 index to capture the U.S. market and STOXX Europe 600 to capture the European market. The S&P 500 index comprises the 500 largest companies in the United States, exclusively consisting of large-cap companies and representing approximately 80% of the market capitalisation<sup>6</sup>. The STOXX Europe 600 index includes the 600 largest companies in Europe, covering large-cap, mid-cap, and small-cap stocks, and accounting for approximately 90% of the European market capitalisation<sup>7</sup>. Both indices serve as robust benchmarks for the empirical analysis due to their broad coverage of diverse stocks.

The observation period spanned over 14 years, starting in 2010 and ending in 2023. This duration incorporates the financial downturn induced by the COVID-19 pandemic and recent regulatory alterations in Europe and, to some extent, in the U.S., enhancing the reliability and availability of ESG data. The deliberate selection of an extended observation period considers the varied outcomes reported in several studies based on the chosen timeframe (De Spiegeleer et al., 2023; Das et al., 2018; Bauer et al., 2005). This

---

<sup>6</sup> S&P 500®. (n.d.). S&P Dow Jones Indices. Retrieved February 6, 2024, from <https://www.spglobal.com/spdji/en/indices/equity/sp-500/#overview>

<sup>7</sup> STOXX® Europe 600. (n.d.). AXIOMA - STOXX - DAX. Retrieved February 4, 2024, from <https://qontigo.com/index/sxxp/>

extended timeframe also enabled the segmentation of the sample into two distinct periods to assess whether there have been any shifts in the overall patterns over time: an early period spanning from 2010 to 2016, inclusive, and a recent period covering the years from 2017 to 2023, inclusive.

This thesis extended its scope to include a more detailed analysis of a downturn period, particularly during the recent COVID-19 pandemic. To achieve this, this paper differentiated between two distinct phases within the pandemic: the drawdown period, characterised by a significant stock market crash resulting in substantial losses of market capitalisation within a short timeframe, and the subsequent recovery period, marked by the stock market rebounding and reaching all-time highs again. The drawdown period is defined from the 20th of February to the 7th of April 2020 (Frazier, 2021) , and the recovery period from the 8th of April to the end of August 2020, coinciding with the time when the MSCI World Index has recovered and reclaimed its all-time high<sup>8</sup>.

This study retrieved the constituents list of both indices from Refinitiv Datastream as of January 1st of each year. For each stock in these lists, data on market capitalisation and the ESGC score (explained in Chapter 5.2.2) as of that date were collected. Additionally, total stock returns were extracted using Refinitiv Datastream's Return Index, which accounts for corporate actions such as dividends and stock splits. This index adds the dividend amount to the closing price on the ex-dividend date. All returns are expressed in U.S. Dollars. Total stock returns were gathered for every month, precisely on the 1st day of each month, resulting in a total of 168 observations over the 14 years. Given the constrained observation window for the COVID-19 analysis, relying on monthly return data became impractical. Consequently, daily return measures were employed for that analysis.

In this context, the monthly return, or the daily returns in the case of the COVID-19 pandemic analysis, of stock  $i$  in month  $t$ , denoted as  $R_{ti}$ , were calculated. The formula for this calculation is expressed as follows:

$$R_{ti} = \frac{RI_{ti} - RI_{(t-1)i}}{RI_{(t-1)i}} \quad (7)$$

$RI_{ti}$  denotes the total return index of stock  $i$  in month (or day)  $t$ .

### 5.2.2 ESG Data – Refinitiv ESGC score

This study collected data from Refinitiv to evaluate companies' sustainability performance, particularly focusing on the ESGC score, which represents the ESG combined score. This metric adjusts the company's basic ESG score derived from

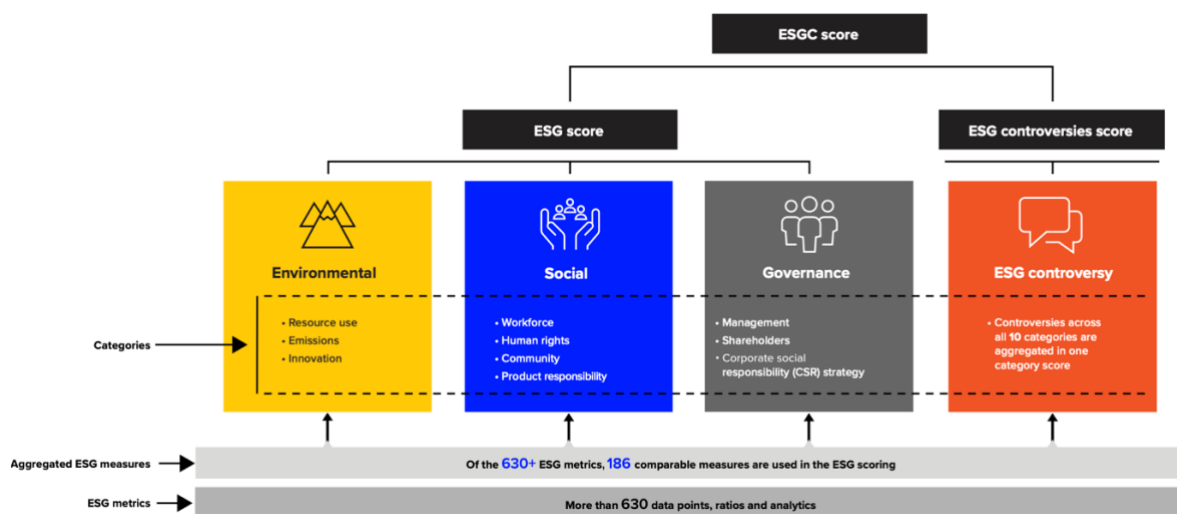
---

<sup>8</sup> Refer to the MSCI website for a visual representation: <https://www.msci.com/end-of-day-history?chart=regional&priceLevel=0&scope=R&style=C&asOf=Jan%2020,%202021&currency=15&size=36&indexId=106>

sustainability reporting by factoring in controversies, thereby providing a more comprehensive understanding of overall sustainability.

Refinitiv offers ESG impact ratings, assessing companies' "relative ESG performance, commitment, and effectiveness"<sup>9</sup>, with a notable advantage being their high transparency in scoring methodology. Unlike certain providers, Refinitiv openly shares its methodology, offering transparency regarding how it calculates ratings. Their extensive ESG database covers over 88% of global market capitalisation and includes data on over 15,000 global companies<sup>9</sup>. As outlined in Figure 1, Refinitiv (2022) collects over 630 ESG metrics for each company from diverse sources like annual reports, company websites, NGO websites, stock exchange filings, CSR reports, news sources, etc. From this dataset, they select the 186 most relevant and comparable metrics, which are grouped into 10 subcategories representing the company's ESG performance. These subcategories contribute to computing Environmental (E), Social (S), and Governance (G) pillar scores. Additionally, they contribute to calculating the overall ESG score, with data aggregated based on various category weights determined by the company's industry and country. Simultaneously, material ESG controversies within these categories, such as lawsuits, ongoing legislative disputes, or fines, are assessed to generate an ESG controversies score. Ultimately, these scores - ESG and ESG controversies - are combined to form the ESGC score.

**Figure 1 : Refinitiv ESG rating Methodology (Source: Refinitiv (2022))**



To ensure objectivity, Refinitiv follows five key principles in its ESG calculations. These principles include:

### 1. Incorporating unique ESG materiality weightings for each industry

Materiality weightings adjust the impact of ESG factors based on their industry-specific relevance, reflecting the varying significance of ESG challenges and risks across sectors.

<sup>9</sup> LSEG. (n.d.). *LSEG ESG scores*. Retrieved November 26, 2023, from <https://www.lseg.com/en/data-analytics/sustainable-finance/esg-scores>

While enhancing the accuracy of ESG scoring methodologies, critics contend that these weightings may oversimplify assessments and overlook company-specific risks (Doyle, 2018).

## **2. Stimulating transparency through company disclosure**

Companies are expected to disclose relevant ESG data and information about their sustainability practices. When calculating a company's ESG score, data points that are not considered highly important for assessing its sustainability have only a minor impact on the overall score. Conversely, the absence of essential and significant ESG data points in a company's reporting adversely impacts its ESG score, thereby fostering company size bias, a prominent limitation associated with ESG scores.

## **3. Establishing industry and country benchmarks**

This entails establishing industry or country-specific benchmarks for ESG performance based on average or best practices. Environmental and social benchmarks are set by industry groups, while governance benchmarks are determined by the country of incorporation. These benchmarks facilitate comparisons of a company's ESG performance with its peers, aiding stakeholders in assessing its sustainability practices relative to competitors. By identifying leaders and laggards in sectors or regions, investors and stakeholders can incentivise ESG improvements, fostering transparency and accountability.

## **4. Utilising a percentile rank scoring methodology**

This provides a standardised and easily interpretable framework for evaluating ESG performance. By assigning scores between 0 and 100, companies' sustainability efforts can be precisely measured and ranked relative to their peers.

## **5. Implementing an ESG controversies overlay to derive the ESGC score**

It verifies companies' actions against their commitments and highlights the influence of significant controversies on the overall ESG score. This ensures a comprehensive evaluation, capturing both positive sustainability efforts and any negative incidents that could affect a company's reputation and long-term sustainability. The overlay also addresses market cap bias by adjusting controversy scores based on company size because larger companies tend to attract more media attention. This adjustment is significant as it seeks to effectively address company size bias in the controversies score. At the same time, a similar correction is not applied to the actual ESG score, as discussed in point 2.

In conclusion, Refinitiv's ESG scores have one notable drawback, namely the potential encouragement of company size bias, whose presence was confirmed by Dobrick et al. (2023). Some researchers suggest addressing this challenge at the ESG score provider level by introducing a 'Missing Pillar Score.' This score would assess companies' potential to disclose missing information, thereby enhancing the overall effectiveness of ESG

evaluations (Sahin et al., 2022). Aside from the company size bias, the overall methodology of Refinitiv's ESG scores is sturdy. While Doyle's (2018) criticism regarding industry materiality may hold some validity, it is important to acknowledge that scoring companies based on individual materiality would be impractical. Thus, employing an industry-based materiality weighing appears to be the most appropriate and feasible approach.

Refinitiv ESG data undergoes annual updates, providing a comprehensive historical perspective on past ESG information. Additionally, this data source is frequently employed by researchers, as evidenced by studies such as Bannier et al. (2019), Amon et al. (2021), Petridis et al. (2023), and Teti et al. (2023).

This study employed the ESG combined score (ESGC score), representing the company's ESG score adjusted in the presence of controversies. Derived from the controversy score, which starts at 100 and decreases when controversies arise, the ESGC score is calculated as the minimum value between the ESG score and the average of the ESG score and the controversy score. This design ensures that the controversy score can only have a detrimental impact on the ESGC score. The scores are measured on a percentile scale ranging from 0 to 100, where 100 signifies the highest attainable score. The rationale behind utilising the ESGC score lies in its comprehensiveness, considering both the theoretical score derived from published data and the real-world factors that may adversely impact it. This score has been employed in studies conducted by Petridis et al. (2023) and Teti et al. (2023).

### **5.2.3 Integration and Refinement of Data Sources**

As discussed in the preceding sections, this study utilised data consisting of the constituents of the S&P 500 and STOXX Europe 600 indices, along with their corresponding Refinitiv ESGC scores. These datasets served as fundamental pillars for subsequent portfolio construction endeavours. Given the annual updates to ESGC scores and the dynamic nature of constituent lists due to various factors such as new additions, removals, insolvencies, mergers, acquisitions, or shifts in market capitalisation, it is important to rebalance portfolios periodically. This highlights the significance of meticulous data integration and cleansing to ensure consistency and reliability. Before portfolios could be constructed, the presence of both return data and ESGC score data for every listed stock across each constituent list had to be confirmed.

Refinitiv releases its ESG ratings retroactively, meaning that the portfolio in year  $t$  must be constructed using ESGC scores from year  $t-1$ , as scores for year  $t$  are not available at the time of the portfolio construction. Therefore, the presence of return data for year  $t$  and ESGC data for year  $t-1$  for each stock in the constituent list at time  $t$  must be confirmed. If observations were missing for ESGC data or return data, often due to events such as mergers, acquisitions, delistings, or insolvencies, those companies were excluded from the available stock universe for that year. However, an exception was made for the year 2023, which needed ESGC scores from 2022. If the ESGC scores from 2022 were missing but

available from 2021, those stocks were not excluded. Instead, it was assumed that the 2022 score would equal the 2021 score. The decision was based on observing a lower level of published 2022 ESGC scores (approximately 30%) compared to ESGC scores from previous years at the time of data extraction<sup>10</sup>. Given this, the probability is high that these scores would have been published in the following months. Omitting these stocks, while unable to omit stocks in previous years where ESGC score publications have been delayed, would have introduced bias to the stock universe.

## 5.2.4 Descriptive Statistics

The post-screening number of constituents fluctuated from 478 to 497 companies annually for the S&P 500 and 576 to 595 companies for the STOXX Europe 600 (see Table 1). In general, earlier years typically saw more excluded stocks compared to recent years due to more missing data points. However, the overall size of the stock universe remained similar across all years.

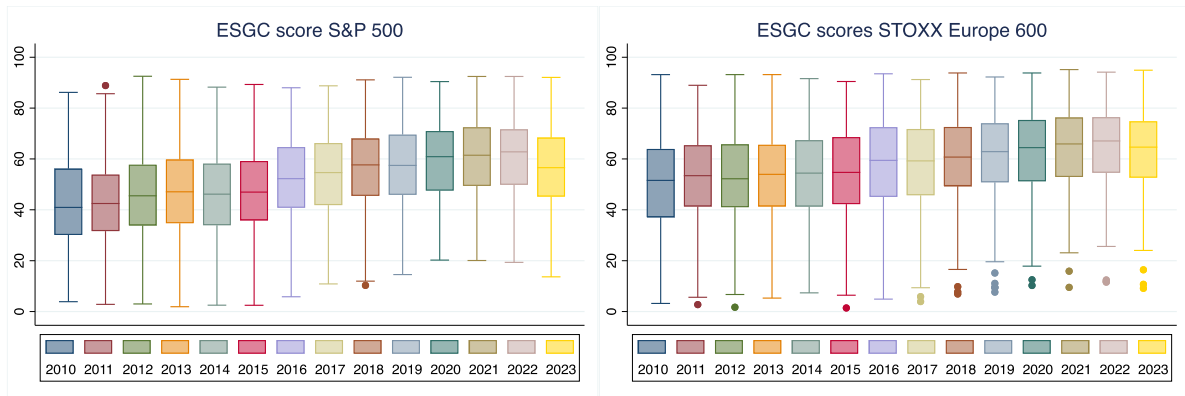
**Table 1** : Pre- and post-screen number of companies and average ESG and ESGC scores across indices and years

Year	S&P 500				STOXX Europe 600			
	#pre-screen	#post-screen	avg. ESG	avg. ESGC	#pre-screen	#post-screen	avg. ESG	avg. ESGC
2010	500	478	48.0	43.4	600	576	53.2	50.2
2011	500	482	49.3	43.3	601	578	55.6	51.9
2012	500	484	51.3	46.2	600	587	56.0	52.0
2013	500	484	51.3	47.3	600	586	56.2	52.6
2014	500	489	51.5	46.4	600	593	56.9	53.3
2015	502	480	52.3	47.7	601	591	57.7	54.4
2016	504	481	55.7	52.7	600	584	59.8	57.7
2017	505	490	57.6	53.6	600	587	61.3	57.9
2018	505	491	59.9	56.7	600	588	64.0	60.3
2019	505	494	61.2	57.5	600	594	66.3	62.0
2020	505	497	63.1	59.0	600	588	68.1	63.0
2021	505	494	65.3	60.4	600	591	69.4	64.0
2022	505	486	66.8	60.3	600	594	69.5	65.1
2023	503	493	67.2	56.7	600	595	69.5	63.1

Moreover, both indices showed an upward trend in average ESG and ESGC scores over time, with STOXX Europe 600 scores significantly surpassing those of the S&P 500, indicative of a more advanced regulatory framework for sustainability reporting, as suggested by Sipiczki (2022). While the average ESG score continued to increase, the average ESGC score has experienced a significant decline since 2022 (reflecting the score of 2021), likely influenced by rising concerns about greenwashing and increased media scrutiny, resulting in more controversies and, subsequently, a reduced ESGC score. Additionally, Figure 2 highlights a narrowing spread of ESGC scores over time, likely influenced by an overall increase in adherence and consistency of ESG reporting practices.

<sup>10</sup> The data was extracted on 05/02/2024

**Figure 2 :** Comparative Boxplots of ESGC scores across indices and years



The boxplots reveal several outliers in the ESGC scores within the STOXX Europe 600 dataset and two outliers within the S&P500 dataset. These outliers, denoted by data points falling 1.5 interquartile ranges below the third quartile or above the first quartile, are considered unusually high or low compared to the rest of the dataset. Initially, the decision was made not to remove these outliers due to their limited number, making it unlikely that they would substantially alter the statistical findings and due to uncertainty regarding their origins. These outliers may stem from various factors, including insufficient reporting, significantly lower sustainability compared to the median company, or a high prevalence of controversies that substantially reduce the overall ESGC score. Omitting these outliers may introduce a bias to the results to the same extent as not excluding them. Nonetheless, as part of the robustness analysis, the outliers were excluded to ensure that the results were not overly dependent on these few stocks. Appendix 7 provides an overview of companies identified as outliers for specific years.

**Table 2 :** ESGC score distribution across indices and years

Year	S&P 500					STOXX Europe 600				
	min	20 perc.	median	80 perc.	max	min	20 perc.	median	80 perc.	max
2010	3.9	27.6	40.9	59.5	86.2	3.2	33.2	51.6	67.0	93.2
2011	2.8	28.7	42.5	57.6	88.9	2.7	37.2	53.4	67.6	89.0
2012	3.0	32.0	45.5	62.7	92.5	1.7	37.8	52.2	68.4	93.2
2013	1.9	32.4	47.1	63.5	91.3	5.3	37.3	53.9	68.9	93.2
2014	2.5	32.0	46.1	61.7	88.2	7.3	37.6	54.5	69.7	91.6
2015	2.5	33.1	47.0	62.0	89.3	1.4	38.5	54.7	71.7	90.5
2016	5.8	38.9	52.3	68.0	88.0	4.9	42.1	59.4	75.2	93.5
2017	10.9	39.2	54.6	68.3	88.8	3.9	42.9	59.3	74.0	91.3
2018	10.3	43.1	57.7	70.7	91.1	6.9	46.3	60.7	75.3	93.8
2019	14.6	43.2	57.5	71.8	92.1	7.6	49.2	62.9	76.9	92.2
2020	20.2	45.0	60.9	72.8	90.4	10.2	48.9	64.5	77.7	93.8
2021	20.1	46.2	61.5	74.4	92.4	9.5	50.2	65.9	78.8	95.2
2022	19.3	46.2	62.8	73.6	92.5	11.6	51.7	67.1	78.6	94.1
2023	13.7	42.5	56.6	70.7	92.1	9.1	50.6	64.6	76.7	94.9

Table 2 underscores the distribution of ESGC scores, revealing a consistent presence of high-performing sustainable companies in both geographic regions, as evidenced by

similar maximum values across the years. However, notable upward shifts are observed in the 20<sup>th</sup>, 50<sup>th</sup>, and 80<sup>th</sup> percentiles, suggesting an overall improvement in sustainability practices or reporting among companies over time, likely influenced by evolving consumer expectations and regulatory standards.

**Table 3** : Average market capitalisation (in million USD) across indices, years and ESGC score

Year	S&P 500			STOXX Europe 600		
	entire Index	bottom 20% ESGC	top 20% ESGC	entire Index	bottom 20% ESGC	top 20% ESGC
2010	21,001	11,155	20,779	15,501	3,762	21,768
2011	24,092	13,429	19,761	15,944	6,262	19,309
2012	23,670	16,940	22,628	13,558	3,796	21,527
2013	26,521	21,731	27,051	15,450	5,787	20,985
2014	34,459	28,416	28,083	19,240	8,537	27,108
2015	38,304	38,587	31,918	17,851	8,716	23,481
2016	37,789	41,867	43,299	17,262	9,286	33,700
2017	39,978	49,876	44,675	16,870	10,839	26,400
2018	47,440	67,877	52,710	20,925	14,634	28,151
2019	43,587	60,406	40,689	17,404	12,905	23,334
2020	55,645	87,406	44,911	20,867	17,883	30,233
2021	66,606	127,594	55,213	22,578	26,425	25,076
2022	85,759	177,262	61,595	25,786	26,878	31,039
2023	67,073	118,969	37,607	21,094	29,855	17,750

Note: This table presents the average market capitalisation (in million USD) per company for each portfolio, including the entire Index, the bottom 20% ESGC, and the top 20% ESGC, for both the S&P 500 and the STOXX Europe 600. Market capitalisation values are as of January 1st of each year.

Table 3 underscores significant differences in average market capitalisation per company in different portfolios. While the average market capitalisation in the S&P 500 increased considerably over the years, the one in the STOXX Europe 600 exhibited only a modest uptrend. Furthermore, variations in average market capitalisation are observed across different ESGC score cut-offs within each index. In contrast to the common small-cap bias found in traditional ESG scores, the ESGC score demonstrates a distinct bias with no clear-cut pattern. In the early years, companies with high ESGC scores had higher market capitalisation than those with lower scores, suggesting a small-cap bias. However, this trend evolved significantly over time, resulting in similar average market capitalisation for both cut-offs in the case of the STOXX Europe 600 index, and considerably higher market capitalisation for the bottom ESGC cut-off in the case of the S&P 500 index. Interestingly, the top ESGC cut-off maintained a similar average market capitalisation compared to the overall index market cap over the observation period, while significant changes were primarily observed in the bottom cut-off. One plausible explanation, consistent with the ESGC score methodology and the traditional small-cap bias, is that the significance of the controversy overlay may have fluctuated over time. Initially, small-cap companies tended to have lower overall ESG scores due to limited sustainability reporting, but these scores gradually increased. Simultaneously, controversies, influenced by factors such as media presence and greenwashing concerns, may have increased significantly over time, resulting in increased scrutiny of sustainability claims, particularly among large-cap companies representing themselves as leaders in

sustainability. As time progressed, this controversy overlay negatively impacted the ESGC scores of high-market capitalisation companies. Questions arise regarding the effectiveness of this overlay, particularly for larger companies like those in the S&P 500. However, the bias in ESGC scores is notably reduced compared to traditional ESG scores, as shown in Appendix 6.

## **5.3 Methodology**

### **5.3.1 Portfolio Formation & Rebalancing**

Each year, on the 1<sup>st</sup> of January, the portfolios underwent rebalancing to accommodate changes in the constituents list and ESGC scores. Other researchers adopted a different approach by utilising the most recent constituent list without modifying the constituents and only adjusting the portfolios based on ESG data. Some studies either considered only those companies of the most recent constituent list for which return and ESG data were available for all years, maintaining a constant sample size (Amon et al., 2021), or gradually increased the sample size as data became available over time (De Spiegeleer et al., 2023). The rationale for annual adjustments to constituents was to maintain a similar sample size over 14 years while avoiding the need to exclude companies entirely due to the unavailability of data during specific periods. This methodology guarantees that the portfolios closely mirror the authentic indices following data cleansing each year and reduces survivorship bias. Consequently, rebalancing occurred annually, exclusively including companies with both ESGC scores and return data available for the entire 12-month observation period within a given year.

Since this study examines whether integrating ESG considerations results in abnormal returns, the portfolios had to be constructed based on published ESGC scores. Hence, in the backtesting process, ESGC scores from period  $t-1$  were used for returns at period  $t$ , similar to Kempf & Osthoff (2007). This practice mirrors the historical reality where investors could only act on the previously published scores, given that the scores for the current year were not available at the time of rebalancing. Using ESGC scores from time  $t$  for return calculations at time  $t$  would essentially be akin to insider trading, as this information was not yet publicly disclosed at that point.

In an initial step, three portfolios for both the S&P 500 and STOXX Europe 600 indices were constructed:

- T20E: Includes companies within the top 20% of the ESGC score distribution
- B20E: Includes companies within the bottom 20% of the ESGC score distribution
- L-S: Long-short hedging strategy, achieved by simultaneously buying T20E and selling B20E

The long-short strategy, also commonly known as a zero-cost investment, involves shorting or selling the low ESG quintile while simultaneously going long or buying the high ESG quintile. This is a commonly employed practice in comparing top portfolios to

bottom portfolios, as seen in studies by Teti et al. (2023), Bannier et al. (2019), Lee et al. (2010), and others.

All three portfolios adopt an equal weighting strategy, denoted by the character "E" in their names. This approach eliminates market capitalisation bias and ensures that portfolio performance is not dictated by the performance of a select few stocks. The weight assigned to each stock is directly proportional to the total number of stocks in the specified portfolio and can be expressed as:

$$w_i = \frac{1}{N} \quad (8)$$

The symbol  $w_i$  represents the weight assigned to asset  $i$ , and  $N$  denotes the total count of stocks included. A market value-weighted approach is discussed later in the analysis.

These portfolios allow an investigation of the general relationship between the ESGC score and risk-adjusted financial performance by comparing the risk-adjusted performance of companies within the top and bottom ESGC score distribution.

In the second phase, four supplementary portfolios were constructed using various weighting methodologies to modify the sustainability attributes of the portfolios without excluding companies. This approach drew inspiration from Amon et al. (2021). Thus, the following portfolios were constructed:

- MKTE: Includes all the companies in the index post-cleansing, with equal weights as per Formula 8.
- MKTV: Includes all the companies in the index post-cleansing, with market value weights. This means that the weight of each stock is contingent on its market value relative to the overall market value and can be articulated as:

$$w_i = \frac{MV_i}{\sum_{k=1}^N MV_k} \quad (9)$$

where  $MV_i$  stands for the market value of company  $i$ .

- ESGE: Includes all the companies in the index post-cleansing, by employing ESGC weighting. This means that the weight assigned to each stock is determined by its ESGC score in relation to the total sum of ESGC scores of all constituents and can be expressed as:

$$w_i = \frac{ESGC_i}{\sum_{k=1}^N ESGC_k} \quad (10)$$

where  $ESGC_i$  stands for the ESGC score of the company  $i$ .

- ESGV: Includes all the companies in the index post-cleansing, with weights determined by a combination of ESGC and market value. In this scenario, the weight assigned to each stock is influenced by both market value and ESGC score and can be formulated as:

$$w_i = \frac{MV_i \cdot ESGC_i}{\sum_{k=1}^N MV_k \cdot ESGC_k} \quad (11)$$

These portfolios allow to evaluate whether integrating ESGC weighting criteria could result in abnormal returns or produce comparable returns while improving the average ESGC score. The idea is to assign higher weights to more sustainable companies without excluding those with lower sustainability scores, thus maintaining a broader diversification pool. Once more, the characters "E" and "V" at the end of the portfolio names signify the weighting approach, where "E" denotes equal weighting and "V" indicates market value weighting. The ESGE portfolio corresponds to the equal-weighted portfolio MKTE when subject to ESGC weighting. Similarly, the ESGV portfolio corresponds to the market value-weighted portfolio MKTV when subject to ESGC weighting. As ESGE and ESGV differ from MKTE and MKTV solely through the addition of an ESGC weight overlay, the performance of these portfolios was assessed using an artificial long-short strategy. In this strategy, a portfolio subject to ESGC-weighting is purchased, while a traditional portfolio is sold simultaneously.

In a later stage, turnover fees were integrated into MKTE, ESGE, MKTV, and ESGV, adopting a methodology akin to Kempf & Osthoff (2007), to assess the net risk-adjusted return that an investor could have achieved by adhering to a more sustainable investment portfolio. Unlike Kempf & Osthoff (2007), only the annual cost of adjusting the portfolio, i.e., the annual turnover costs, was considered, excluding the costs associated with opening or closing the portfolio, as these costs could be deemed identical for both portfolios. Following a similar method inspired by Derwall et al. (2005), round-trip transaction costs ranging between 50 and 200 basis points were assumed, as calculating specific transaction costs for each portfolio was beyond the scope of this study. Turnover is defined as the total value of transactions, which includes both purchases and sales, required for rebalancing the portfolio on the 1st of January. This turnover is expressed as a ratio relative to the portfolio's total assets on that specific date. The detailed methodology for calculating the turnover rate is provided in Appendix 5. Transaction costs were incorporated by multiplying the annual turnover rate by the round-trip transaction costs, which were subsequently subtracted from the portfolio's return on the 1st of January each year. The portfolios, inclusive of turnover costs, were also compared using an artificial long-short strategy, wherein the performance of one portfolio, after considering costs, was compared to the performance of another portfolio after accounting for costs.

### 5.3.2 Performance Measurement

The financial performance of the various portfolios was calculated through the portfolio returns. These returns were calculated by multiplying the return of each stock (as calculated by Formula 7) by its weight in the portfolio:

$$R_t = \sum_{i=1}^N w_i \cdot R_{ti} \quad (12)$$

Where  $R_t$  signifies the portfolio return in month  $t$ ,  $w_i$  represents the weight of stock  $i$  in the portfolio,  $R_{ti}$  denotes the return of stock  $i$  in month  $t$ , and  $N$  represents the total number of stocks in the portfolio.

Risk-adjusted returns were calculated via the Sharpe ratio during the initial descriptive statistics. The Sharpe ratio quantifies a portfolio's excess return per unit of risk. A higher Sharpe ratio indicates a better risk-adjusted performance. Investors and fund managers often use this ratio to compare the risk-adjusted returns of different investments or portfolios. The formula for the Sharpe ratio is:

$$\text{Sharpe Ratio} = \frac{R_p - r_f}{\sigma} \quad (13)$$

In this context,  $R_p$  represents the anticipated portfolio return,  $r_f$  denotes the risk-free rate of return, and  $\sigma$  signifies the standard deviation of the portfolio's excess return - defined as the difference between the portfolio return and the risk-free rate. The expected portfolio return was calculated through the annualised geometric average historical return. The formula is the following:

$$R_p = \left[ \prod_{t=1}^T (1 + R_t) \right]^{\frac{12}{T}} - 1 \quad (14)$$

Where  $R_t$  represents the portfolio's return in month  $t$  (as in Formula 12), and  $T$  represents the total number of months. Likewise, the average risk-free rate was annualised to obtain the required  $r_f$  for the Sharpe ratio calculation. The monthly risk-free rate was sourced from Kenneth R. French Data Library and is based on the 1-month T-Bill rate<sup>11</sup>.

The annualised standard deviation of the portfolio's return is determined by taking the square root of the annualised variance, which is the average of the squared difference between each month's return and the average return, following this formula<sup>12</sup>:

---

<sup>11</sup> *Kenneth R. French - Data Library*. (n.d.).

[http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data\\_library.html](http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html)

<sup>12</sup> In computing the Sharpe Ratio, the annualised standard deviation of the excess return, i.e.,  $R_p - r_f$ , is employed

$$\sigma_p = \sqrt{12 \frac{\sum_{t=1}^T [R_t - \bar{R}]^2}{T - 1}} \quad (15)$$

$\bar{R} = \frac{1}{T} \sum_{k=1}^T R_k$  is the arithmetic average return of the portfolio.

Amidst the COVID-19 pandemic, marked by a stock market crash and increased volatility, it's essential to examine supplementary ratios. These ratios, such as maximum daily loss, maximum drawdown, and the Sortino ratio, aid in quantifying downside risk and potential investment losses. Each metric provides additional insights into the portfolios' performance and risk profiles during the specified periods.

The Sortino ratio, an enhanced iteration of the Sharpe ratio, differs by utilising downside deviation as a measure of risk instead of standard deviation. This deviation is based on the principle that investors should primarily focus on downside risk. Upside deviation accounted for in the Sharpe ratio computation cannot be considered as risk and should thus be excluded. A heightened Sortino ratio indicates a superior return per unit of downside risk. Rollinger & Hoffman (2013) defines the Sortino ratio as:

$$\text{Sortino Ratio} = \frac{R_p - TR}{TDD} \quad (16)$$

$R_p$  denotes the portfolio return, following the definition in Formula 14, and  $TR$  stands for the target return, which is 0 for this study due to risk-free rates being close to zero during the COVID-19 pandemic.  $TDD$ , the target downside deviation, is defined as the root-mean-square of the deviations of the realised returns' underperformance from the target return. Mathematically, this involves treating all returns above the target return as underperformance with a value of 0.

This concept can be expressed as:

$$TDD = \sqrt{\frac{1}{T} \sum_{t=1}^T [\text{Min}(0, R_t - TR)]^2} \quad (17)$$

Where:

$R_t$  = return on day  $t$

$T$  = total number of returns

$TR$  = target return, for this specific study 0

Maximum drawdown captures the most substantial decline in a portfolio's value from the peak to the trough over a defined period. It sheds light on potential losses that could transpire during challenging market conditions (Magdon-Ismail & Atiya, 2004). Finally, the maximum daily loss defines the portfolio's most significant decline within a single day.

### 5.3.3 Multiple Linear Time Series Regression

Multi-factor regression models were employed to investigate whether the various ESG portfolios developed in this study can yield abnormal returns. This methodology is widely used by researchers, and numerous prior studies have employed some variant of multiple linear time series regression. The multiple regression model assesses the influence on the dependent variable when adjusting one of the independent variables while keeping the other regressors constant.

Most studies in this field employed either the Carhart 4-factor model (Kempf & Osthoff, 2007; Statman & Glushkov, 2009; Halbritter & Dorfleitner, 2015; Bannier et al., 2019) or the Fama-French 5-factor model (Amon et al., 2021; Milonas et al., 2022; Teti et al., 2023). As discussed in Chapter 3.5, both models are generally regarded as the most effective in explaining portfolio returns. This study opted for the Fama-French 5-factor model, considering it is the latest multi-factor model published and has been employed in the most recent studies. However, to ensure the robustness of the results and mitigate potential dependencies on model specifications, this paper conducted an additional set of time-series regressions using the Carhart 4-factor model in the robustness analysis.

The formula for the Fama-French 5-factor model, outlined in Chapter 3.5, is expressed as follows:

$$R_{it} - R_{ft} = \alpha_i + \beta_1(R_{Mt} - R_{ft}) + \beta_2SMB_t + \beta_3HML_t + \beta_4RMW_t + \beta_5CMA_t + \varepsilon_{it}$$

This model controls for the following factors<sup>13</sup> :

- **R<sub>Mt</sub> - R<sub>ft</sub>** denotes the market's excess return.
- **SMB** quantifies the size factor, revealing the excess return obtained by small-cap stocks relative to large-cap stocks.
- **HML** represents the value factor, showcasing the excess return achieved by value stocks compared to growth stocks.
- **RMW** illustrates the profitability factor, indicating the excess return earned by stocks with strong operating profitability compared to those with weak operating profitability.
- **CMA** represents the investment factor, reflecting the excess return earned by stocks with conservative investment strategies compared to those with aggressive investment strategies.

This model tests whether the portfolio's excess return can be ascribed to the risk factors represented by the independent variables or if it generates abnormal returns, as indicated by the alpha ( $\alpha$ ). Significantly positive alphas from ESG portfolios present an opportunity for investors to capitalise on, while significantly negative alphas provide the potential for abnormal returns through a short-selling strategy.

---

<sup>13</sup> For a more detailed explanation of the different risk factors, please refer to Chapter 3.5

Data needed for this multi-factor model, including market ( $R_{Mt}-R_{ft}$ ), size (SMB), value (HML), profitability (RMW) and investment (CMA), were collected from the Kenneth R. French Data Library<sup>14</sup>. The factors were collected for the U.S. market to represent the dataset on the S&P 500 and for the European<sup>15</sup> market to represent the dataset on the STOXX Europe 600. Due to the difference in calculation periods (monthly stock returns from the first to the first of the month in this study, while Kenneth R. French uses last of month to last of month), precalculated monthly returns were not used. Instead, they were computed based on the daily returns provided

This thesis used Ordinary Least Squares (OLS) regression to scrutinise the dependent variable, represented by the excess returns of the portfolios, in relation to the independent variables that include the five factors used in the Fama-French 5-factor model. According to Stock & Watson (2020), the effectiveness of multi-factor time series regression using OLS relies on crucial assumptions. These include verifying that the time series is stationary, maintaining constant statistical properties such as mean and variance over time; homoscedastic, with consistent variance in the residuals of the regression; and free from autocorrelation, ensuring no correlation among the residuals of the regression model. Additionally, in utilising a multi-factor model, the independent variables mustn't be susceptible to multicollinearity. This condition can be confirmed, as researchers globally have widely researched and adopted the Fama-French model.

These assumptions are crucial and necessitate rigorous examination. In this context, all the essential tests were performed, including the Augmented Dickey-Fuller test for stationarity, White's general test to scrutinise the presence of heteroscedasticity, and the Durbin-Watson test to assess autocorrelation in the residuals. All statistical analyses and regressions were conducted using the STATA statistical software. The results conclude that all portfolios are stationary (Appendix 1.1), and a few individual portfolios exhibit some level of positive or negative autocorrelation (Appendix 1.2). Additionally, most portfolios display indications of heteroscedasticity (Appendix 1.3). To address the adverse effects of heteroscedasticity and autocorrelation, this paper calculated Newey-West standard errors, which correct for heteroscedastic and autocorrelated coefficients estimated by OLS regression. Heteroscedasticity and autocorrelation consistent (HAC) robust standard errors were applied for each portfolio to ensure consistency with the calculated standard errors. In STATA, a truncation parameter, hereinafter  $m$ , is required to determine the maximum lag order of autocorrelation. This study adopted the rule of thumb proposed by Stock & Watson (2020), derived from the following formula:

$$m = \text{int}\left(0.75 \cdot N^{\frac{1}{3}}\right) \quad (18)$$

$N$  represents the sample size, and  $\text{int}(x)$  denotes the integer obtained by truncating  $x$  toward 0.

---

<sup>14</sup> Kenneth R. French - Data Library. (n.d.).

[http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data\\_library.html](http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html)

<sup>15</sup> Includes stocks from Austria, Belgium, Switzerland, Germany, Denmark, Spain, Finland, France, Great-Britain, Greece, Ireland, Italy, Netherlands, Norway, Portugal and Sweden

## 6 Empirical Analysis

This chapter presents the analysis of the various sustainable and non-sustainable portfolios. This section is subdivided into two distinct subsections, each dedicated to specific geographical regions to ensure a comprehensive analysis. Chapter 6.1 outlines the findings related to the United States of America, while Chapter 6.2 details the outcomes observed in Europe.

Within each geographical region, the initial analysis concentrated on the overall performance of portfolios based on top and bottom ESGC scores. Subsequently, a more detailed examination explored their performance during the specific subperiod of the COVID-19 pandemic. Finally, the study shifted its focus to ESGC-weighted portfolios. The analysis encompassed scenarios both with and without the inclusion of transaction costs, aiming to offer investors a more grounded perspective on potential risk-adjusted returns.

The analysis encompassed descriptive elements, providing insights into each portfolio's return, risk, and risk-adjusted performance. Return metrics, such as annualised and cumulative returns, were examined, with risk being evaluated through the lens of annualised standard deviation. The Sharpe ratio was used to measure risk-adjusted returns. Additionally, the analysis extended to results derived from the Fama-French 5-factor regression. The coefficients were presented with four decimal points, and significance levels were examined at 10%, 5%, and 1% confidence intervals. The study commented on significant Fama-French risk factors, with discussions extending to alpha coefficients, even when statistical significance was not apparent. It's important to acknowledge that interpretations drawn from nonsignificant alphas lack robustness and may be influenced by random chance.

Finally, a robustness analysis was conducted to examine how the results react to methodological setup changes and highlight their corresponding impacts. Initially, the study explored a shift from equal-weighted to value-weighted top and bottom portfolios, referred to as T20V and B20V. Subsequently, all regressions in this analysis were tested using the Carhart 4-factor model to assess whether the model specification produced different alpha outcomes. Lastly, an examination was conducted to investigate whether excluding statistical outliers resulted in divergent findings.

### 6.1 United States of America

#### 6.1.1 Top/Bottom Portfolio Analysis (U.S.)

In the analysis of the U.S. market, represented by the S&P 500 index constituents, the equal-weighted T20E and B20E portfolios are compared, using the equal-weighted MKTE portfolio as a benchmark. Table 4 provides performance statistics, indicating that over the entire observation period, T20E outperformed with higher annualised returns (12.94%) compared to B20E (11.37%) and MKTE (12.66%). T20E also exhibited lower

volatility, resulting in a higher Sharpe ratio (70.5%) compared to B20E (55.4%) and MKTE (65.7%).

**Figure 3 :** Cumulative Returns for equal-weighted T20E, B20E and MKTE portfolios based on S&P 500 constituents

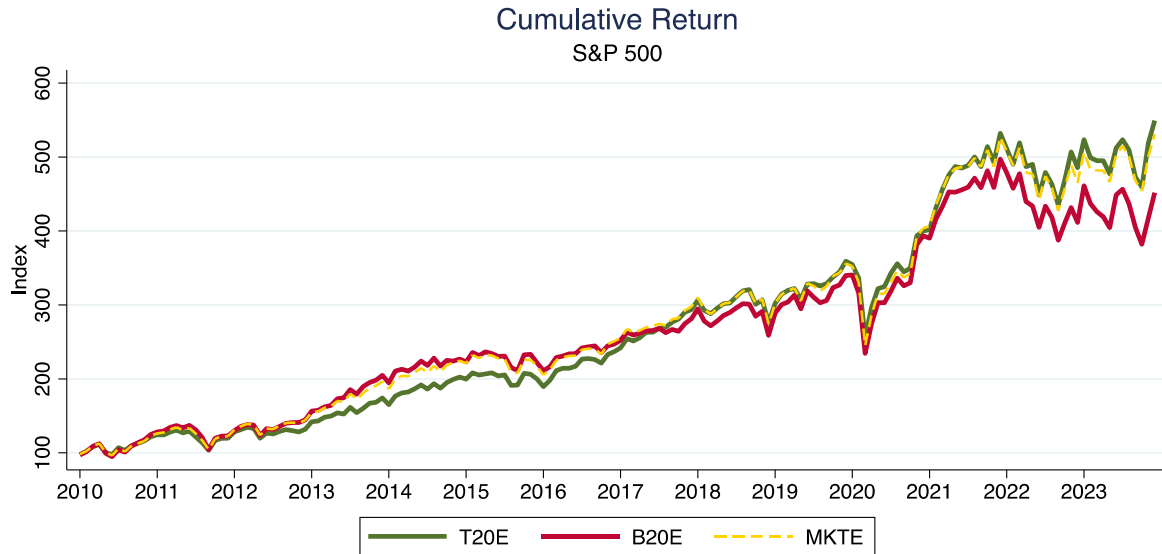


Figure 3 illustrates a shift in performance dynamics, indicating that B20E outperformed until early 2017, and subsequently, T20E demonstrated superior cumulative returns. This underscores the significance of considering distinct time periods in the analysis. Additionally, a significant divergence in cumulative returns from 2021 onwards can be observed. During this period, the B20E portfolio yielded no returns, maintaining a flat curve between the beginning of 2021 and the end of 2023, while the T20E portfolio experienced a significant increase in return over the same period.

**Table 4 :** Performance statistic measures for equal-weighted T20E, B20E and MKTE portfolios based on S&P 500 constituents

	2010-2023			2010-2016			2017-2023		
	T20E	B20E	MKTE	T20E	B20E	MKTE	T20E	B20E	MKTE
Annualised Return	12.94%	11.37%	12.66%	12.84%	13.63%	13.81%	12.77%	9.00%	11.31%
Annualised Volatility	17.13%	18.98%	17.97%	14.66%	16.44%	15.63%	19.37%	21.32%	20.14%
Sharpe Ratio	70.5%	55.4%	65.7%	87.2%	82.6%	88.0%	57.7%	34.7%	48.2%

This shift in performance dynamics is also evident in Table 4. All portfolios showed similar Sharpe ratios during the initial observation period from 2010 to 2016. Although B20E had slightly higher returns than T20E, the increased volatility resulted in lower risk-adjusted returns. On the other hand, T20E, with slightly lower returns but lower volatility, achieved a slightly superior Sharpe Ratio of 87.2%, compared to 82.6% for B20E. However, during the second observation period from 2017 to 2023, a significant divergence in risk-adjusted returns as measured by the Sharpe Ratio became apparent. The T20E portfolio exhibited substantial outperformance with a return of 12.77%, surpassing both MKTE with 11.31% and B20E with a modest 9% return. Furthermore, the B20E portfolio

demonstrated higher volatility, leading to a Sharpe Ratio of only 34.7%, compared to 57.7% for T20E and 48.2% for MKTE.

**Table 5 : Fama-French 5-factor loading on equal-weighted T20E, B20E and L-S portfolios based on S&P 500 constituents**

	2010 - 2023			2010 - 2016			2017 - 2023		
	T20E	B20E	L-S	T20E	B20E	L-S	T20E	B20E	L-S
RMt-Rft	0.9683*** (0.0252)	1.0378*** (0.0220)	-0.0695*** (0.0223)	0.9800*** (0.0279)	1.0522*** (0.0381)	-0.0721** (0.0331)	0.9805*** (0.0357)	1.0340*** (0.0294)	-0.0535* (0.0299)
SMB	0.0777** (0.0383)	0.2052*** (0.0628)	-0.1275*** (0.0486)	0.0423 (0.0431)	0.1056* (0.0541)	-0.0634 (0.0626)	0.0752 (0.0645)	0.2632** (0.1067)	-0.1880** (0.0791)
HML	0.1579*** (0.0568)	0.1256** (0.0540)	0.0323 (0.0441)	-0.0393 (0.0499)	-0.0666 (0.0722)	0.0273 (0.0895)	0.2246*** (0.0701)	0.1537** (0.0694)	0.0709 (0.0613)
RMW	0.1051 (0.0716)	0.0651 (0.0543)	0.0400 (0.0561)	0.1193 (0.0828)	-0.0950 (0.0824)	0.2143** (0.0974)	0.0470 (0.0815)	0.1051 (0.0726)	-0.0581 (0.0745)
CMA	0.1195 (0.0833)	0.0340 (0.0846)	0.0855 (0.0682)	0.1490 (0.1031)	0.0478 (0.0848)	0.1013 (0.1452)	0.1071 (0.1078)	0.0509 (0.1157)	0.0562 (0.0771)
Alpha	0.0002 (0.0009)	-0.0013 (0.0011)	0.0015 (0.0010)	-0.0001 (0.0010)	0.0001 (0.0011)	-0.0002 (0.0014)	0.0009 (0.0013)	-0.0023 (0.0018)	0.0032** (0.0014)
Observations	168	168	168	84	84	84	84	84	84
Adj. R <sup>2</sup>	0.9500	0.9415	0.1941	0.9628	0.9596	0.2139	0.9500	0.9371	0.1845

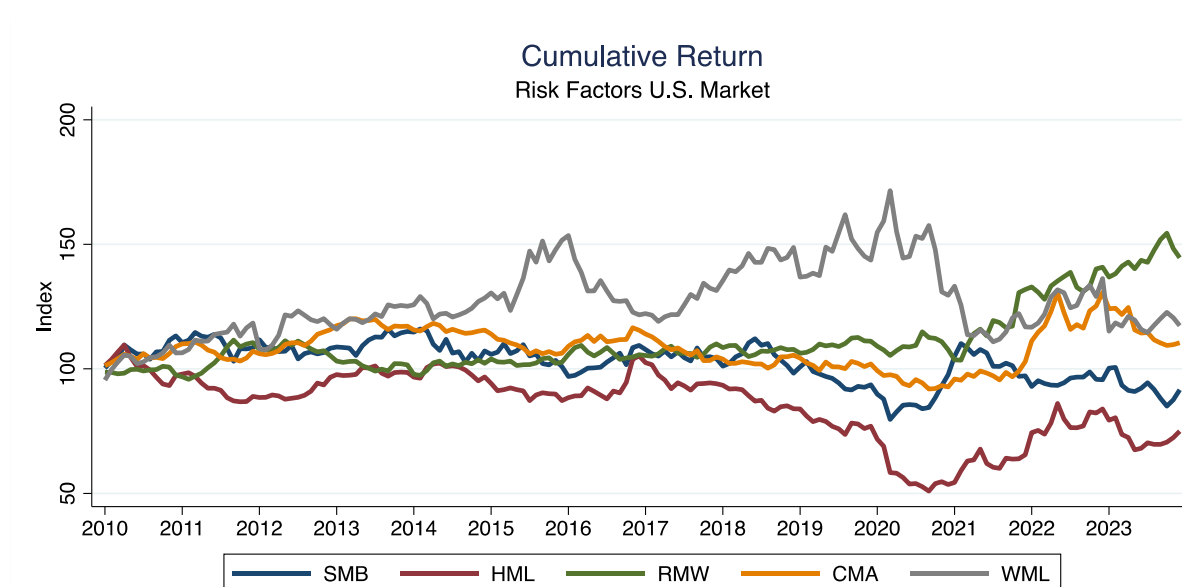
Note: This table illustrates the Fama-French 5-factor loading on the equal-weighted T20E, B20E, and L-S Portfolios, constructed using constituents from the S&P 500 as the underlying basis. The data is categorised into three distinct time periods: 2010-2023, 2010-2016, and 2017-2023. HAC robust standard errors are presented in parentheses. Levels of statistical significance are denoted by \*\*\* (1%), \*\* (5%), and \* (10%).

To assess whether variations in financial performance can be attributed to the sustainability factor or typical risk factor exposures, this study employed the Fama-French 5-factor regression model. This model analyses the excess returns of both portfolios and their respective long-short portfolio after controlling for the SMB, HML, RMW and CMA risk factors. Table 5 presents the regression results, revealing nonsignificant alphas for all portfolios, except the long-short portfolio from 2017-2023, which exhibited a significant positive monthly alpha of 0.32% at the 5% confidence level. This alpha can be explained through a major negative alpha for the bottom portfolio and a smaller positive alpha for the top portfolio, although individual alphas were not statistically significant. Hence, even when considering established risk factors, a positive alpha persisted in the recent period, suggesting a potential deviation from efficient markets and the potential for investors to have achieved abnormal monthly returns of 0.32% by adopting a strategy of buying T20E and selling B20E in recent years. In contrast, the absence of abnormal returns in the period from 2010 to 2016 indicates market efficiency. Over the entire observation period, differences in risk-adjusted returns between T20E and B20E were no longer statistically significant.

Examining the impact of various risk factors on portfolio returns reveals that the market risk factor significantly influenced the excess returns of both portfolios. In contrast, the long-short portfolio maintained a market risk-neutral stance, thanks to its inherent hedge against market risk by simultaneously buying and selling a diversified portfolio. The B20E portfolio exhibited a higher market factor loading than T20E, signifying a greater level of systematic risk, consistent with its higher annualised volatility. Additionally,

B20E exhibited a significant positive exposure to the SMB factor, suggesting a higher allocation to small stocks in its portfolio, thus rendering its returns more sensitive to the performance of these smaller companies. Interestingly, T20E also displayed a positive loading on the SMB factor, though to a lesser and less significant extent. Moreover, both high and low ESGC portfolios exhibited an increased allocation to value stocks in the recent period. Most other risk factors remained statistically insignificant, suggesting that overall, both portfolios possess similar style characteristics, except market capitalisation.

**Figure 4 :** Cumulative Returns for Fama-French & Carhart risk factors in the U.S. market



To further understand how individual risk factors impact risk-adjusted returns, Figure 4 provides a visual representation of the cumulative returns of these risk factors over time. No major trends could be observed in the pre-2017 period. However, in the post-2017 period, growth and large-cap stocks outperformed until the COVID-19 pandemic. Subsequently, in the post-COVID-19 crisis phase (since 2021), value stocks with robust operating characteristics and conservative investment attributes outperformed their peers.

In the context of the T20E and B20E portfolios, the primary divergence in style tilts centred around the SMB factor. However, this factor didn't result in major different returns over the entire period, thus failing to account for return variations. Consequently, the significant positive alpha observed for the long-short strategy since 2017 cannot be attributed to these risk factor differences. This suggests that sustainability has proven financially rewarding in terms of risk-adjusted returns during the recent period.

### 6.1.2 COVID-19 Period (U.S.)

Given the recent subperiod includes a significant financial crisis, it is important to examine the performance of the top and bottom portfolios during the drawdown and recovery periods of the COVID-19 pandemic. Table 6 provides descriptive performance statistics. During the drawdown period, T20E exhibited lower downside risk and superior

risk-adjusted returns than B20E, evidenced by a higher Sharpe and Sortino ratio and a lower maximum drawdown. However, in this scenario, both T20E and B20E underperformed the benchmark. In the recovery period, T20E continued to exhibit lower downside risk, reflected in reduced average downside volatility and maximum drawdown. However, it delivered lower returns, leading to lower Sharpe and Sortino ratios.

**Table 6 :** Performance statistic measures for equal-weighted T20E, B20E and MKTE portfolios based on S&P 500 constituents during the COVID-19 pandemic

	Entire period			Drawdown period			Recovery period		
	T20E	B20E	MKTE	T20E	B20E	MKTE	T20E	B20E	MKTE
Average return	-0,02%	-0,03%	0,01%	-0,80%	-0,94%	-0,58%	0,21%	0,24%	0,24%
Average volatility	3,08%	3,13%	3,21%	5,31%	5,26%	5,56%	1,80%	1,91%	1,88%
Average downside volatility	2,25%	2,28%	2,34%	3,99%	4,01%	4,13%	1,18%	1,23%	1,26%
Highest daily loss	-11,99%	-11,95%	-12,71%	-11,99%	-11,95%	-12,71%	-6,50%	-6,75%	-6,87%
Maximum drawdown	-38,07%	-40,18%	-37,64%	-38,07%	-40,18%	-37,64%	-10,88%	-13,60%	-9,68%
Sharpe Ratio	-0,56%	-0,86%	0,38%	-15,12%	-17,86%	-10,46%	11,45%	12,52%	12,62%
Sortino Ratio	-0,70%	-1,12%	0,57%	-20,01%	-23,31%	-13,97%	17,38%	19,52%	18,79%

Note: The term "Drawdown period" denotes the timeframe from 20/02/2020 to 07/04/2020, while the "Recovery period" spans from 08/04/2020 to 31/08/2020. The "Entire period" encompasses both the Drawdown and Recovery periods.

Table 7 unveils the findings from the Fama-French 5-factor regression. No statistically significant alphas emerged. During the drawdown phase, T20E displayed a higher daily alpha of 0.14%, surpassing B20E's 0.02%.

**Table 7 :** Fama-French 5-factor loading on equal-weighted T20E, B20E and L-S portfolios based on S&P 500 constituents during the COVID-19 pandemic

	Entire period			Drawdown period			Recovery period		
	T20E	B20E	L-S	T20E	B20E	L-S	T20E	B20E	L-S
RMt-Rft	0.9981*** (0.0164)	0.9751*** (0.0208)	0.0230* (0.0123)	0.9928*** (0.0157)	0.9657*** (0.0180)	0.0271 (0.0185)	0.9885*** (0.0380)	0.9876*** (0.0255)	0.0009 (0.0268)
SMB	0.0909 (0.0609)	0.1784*** (0.0595)	-0.0875** (0.0362)	0.0755 (0.1066)	0.1428* (0.0762)	-0.0674 (0.0938)	0.1365*** (0.0515)	0.2275*** (0.0552)	-0.0910* (0.0506)
HML	0.2903*** (0.0470)	0.2849*** (0.0443)	0.0054 (0.0323)	0.3430*** (0.1100)	0.3050** (0.1304)	0.0381 (0.0815)	0.2581*** (0.0396)	0.2674*** (0.0378)	-0.0093 (0.0333)
RMW	-0.1618* (0.0843)	0.2112** (0.0844)	-0.3731*** (0.0579)	-0.1574 (0.2168)	0.3036 (0.2008)	-0.4610*** (0.1184)	-0.1231* (0.0664)	0.1842*** (0.0642)	-0.3073*** (0.0692)
CMA	0.3611*** (0.1180)	-0.0216 (0.1086)	0.3827*** (0.0886)	0.3155 (0.2167)	-0.0460 (0.2036)	0.3615* (0.1964)	0.3742*** (0.1399)	-0.0354 (0.1200)	0.4095*** (0.0933)
Alpha	0.0002 (0.0004)	-0.0000 (0.0004)	0.0003 (0.0003)	0.0014 (0.0014)	0.0002 (0.0017)	0.0013 (0.0010)	-0.0002 (0.0003)	-0.0002 (0.0003)	0.0000 (0.0003)
Observations	135	135	135	34	34	34	101	101	101
Adj. R <sup>2</sup>	0.9825	0.9811	0.3113	0.9898	0.9839	0.2516	0.9586	0.9688	0.2994

Note: This table illustrates the Fama-French 5-factor loading on the equal-weighted T20E, B20E, and L-S Portfolios, constructed using constituents from the S&P 500 as the underlying basis. The data is categorised into three distinct time periods: "Entire period" 20/02/2020-31/08/2020, "Drawdown period" 20/02/2020 – 07/04/2020, and "Recovery period" 08/04/2020 – 31/08/2020. HAC robust standard errors are presented in parentheses. Levels of statistical significance are denoted by \*\*\* (1%), \*\* (5%), and \* (10%).

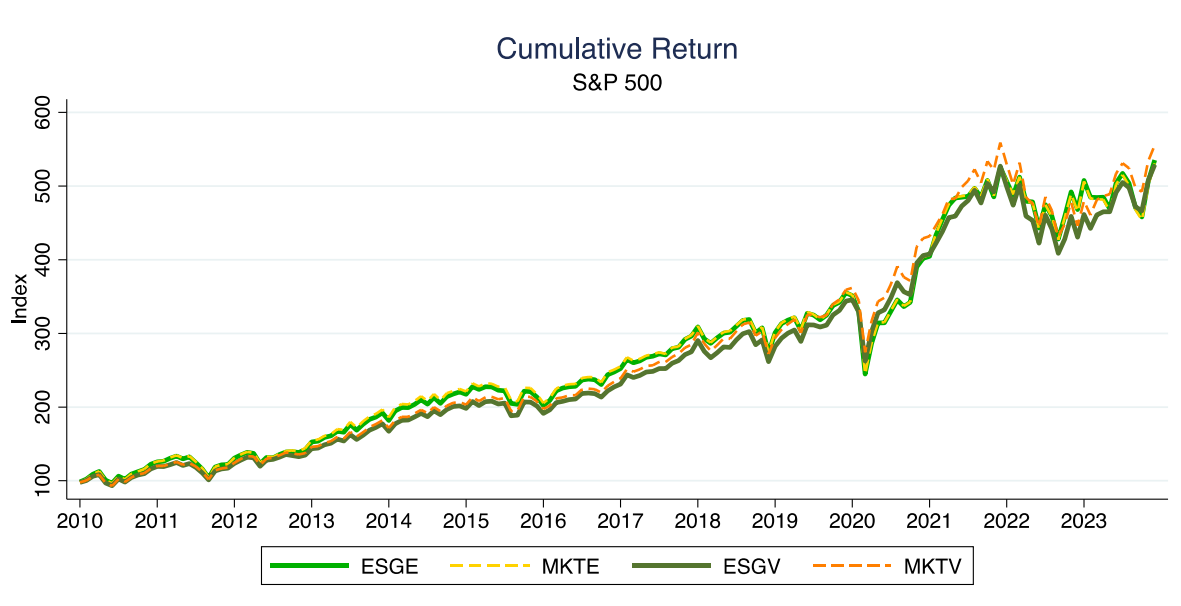
This suggests, in line with the descriptive statistics, that the sustainable portfolio outperformed during the drawdown period, though not to a statistically significant extent. This resulted in a cumulative outperformance of around 4% in this short period.

In the recovery phase, no differences in risk-adjusted returns were observed. Previously identified differences in the descriptive statistics could be explained by loadings on different Fama-French risk factors, specifically RMW and CMA. In 2020, the sustainable portfolio favoured companies with weak operating profitability and a conservative investment strategy. The findings suggest that the sustainable portfolio might tend to outperform during drawdowns but not during recovery periods.

### 6.1.3 ESG-weighted Portfolio Analysis (U.S.)

In the final analysis, this paper examined four portfolios based on different asset allocation strategies. Figure 5 illustrates their cumulative returns, revealing no significant distinctions. The returns of the ESGC-weighted portfolio (ESGE) closely mirrored those of the equal-weighted market portfolio (MKTE), and those from the ESGC and value-weighted portfolio (ESGV) imitated those from the value-weighted market portfolio (MKTV).

**Figure 5 :** Cumulative Returns for ESGE, MKTE, ESGV and MKTV portfolios based on S&P 500 constituents



Descriptive performance statistics in Table 8 further demonstrate a similar alignment regarding risk and reward. Each sustainable portfolio closely mirrored its benchmark when considering the Sharpe ratio. ESGE showed a slight outperformance compared to MKTE, while ESGV exhibited a small underperformance relative to MKTV. Interestingly, value-weighted indexes, favouring large-cap companies, displayed lower volatility compared to equal-weighted indexes. Larger companies, being more mature and less susceptible to systemic risk factors and liquidity fluctuations, generally manifest lower volatility. For instance, small-cap companies are often more sensitive to shifts in monetary policy, particularly concerning changing credit conditions (Conover et al., 2005).

**Table 8 : Performance statistic measures for ESGE, MKTE, ESGV and MKTV portfolios based on S&P 500 constituents**

	2010-2023				2010-2016				2017-2023			
	ESGE	MKTE	ESGV	MKTV	ESGE	MKTE	ESGV	MKTV	ESGE	MKTE	ESGV	MKTV
Annualised Return	12.7%	12.7%	12.6%	13.0%	13.6%	13.8%	12.1%	12.6%	11.6%	11.3%	12.8%	13.1%
Annualised Volatility	17.8%	18.0%	16.2%	16.4%	15.4%	15.6%	14.2%	14.3%	20.0%	20.1%	18.1%	18.3%
Sharpe Ratio	66.8%	65.7%	72.6%	74.4%	88.1%	88.0%	84.4%	87.7%	50.3%	48.2%	62.1%	63.0%

Although only marginal differences in risk-adjusted returns were observed, incorporating ESGC criteria into the portfolio allocation significantly enhanced the overall sustainability profile. Implementing ESGC-weighting in the portfolios led to an increase of approximately 5 additional ESGC points compared to their benchmarks (Appendix 4).

To examine the effect of integrating ESGC weighting into portfolio allocation on returns compared to traditional portfolios while controlling for well-established risk factors, the study conducted Fama-French 5-factor regression on a long ESGE short MKTE Portfolio, as well as a long ESGV and short MKTV portfolio, across different time periods.

**Table 9 : Fama-French 5-factor loading on ESGE – MKTE and ESGV – MKTV difference portfolios using S&P 500 constituents**

	2010 - 2023		2010 - 2016		2017 - 2023	
	ESGE - MKTE	ESGV - MKTV	ESGE - MKTE	ESGV - MKTV	ESGE - MKTE	ESGV - MKTV
RMt-Rft	-0.0078*** (0.0028)	-0.0099** (0.0045)	-0.0101** (0.0042)	-0.0007 (0.0041)	-0.0044 (0.0027)	-0.0151** (0.0063)
SMB	-0.0147*** (0.0049)	0.0080 (0.0065)	-0.0117 (0.0072)	-0.0051 (0.0066)	-0.0195*** (0.0063)	0.0168 (0.0117)
HML	0.0018 (0.0048)	0.0058 (0.0081)	-0.0058 (0.0103)	-0.0136 (0.0109)	0.0080 (0.0055)	0.0113 (0.0108)
RMW	0.0093 (0.0059)	0.0051 (0.0120)	0.0290** (0.0112)	0.0071 (0.0104)	-0.0045 (0.0051)	0.0040 (0.0160)
CMA	0.0074 (0.0074)	0.0381*** (0.0139)	0.0202 (0.0182)	0.0538*** (0.0165)	0.0025 (0.0074)	0.0309 (0.0190)
Alpha	0.0001 (0.0001)	-0.0002* (0.0001)	-0.0001 (0.0002)	-0.0004** (0.0002)	0.0003** (0.0001)	-0.0001 (0.0002)
Observations	168	168	84	84	84	84
Adj. R <sup>2</sup>	0.2526	0.2317	0.3560	0.1121	0.2156	0.3052

Note: This table illustrates the Fama-French 5-factor loading on the difference portfolios ESGE – MKTE (formed by buying ESGE and selling MKTE) and ESGV – MKTV (formed by buying ESGV and selling MKTV). These portfolios are constructed using constituents from the S&P 500 as the underlying basis. The data is categorised into three distinct time periods: 2010-2023, 2010-2016, and 2017-2023. HAC robust standard errors are presented in parentheses. Levels of statistical significance are denoted by \*\*\* (1%), \*\* (5%), and \* (10%).

The results in Table 9 paint a nuanced picture. While incorporating ESGC weighting into equal-weighted portfolios generated a positive alpha in the recent subperiod, aligning with the observations in Chapter 6.1.1, the integration into value-weighted portfolios produced negative alphas, particularly in the earlier subperiod. This unexpected finding is further examined and discussed later. When considering the entire period, only a small statistically insignificant alpha of 0.01% is observed when incorporating ESGC weighting into the equal-weighted portfolio. In comparison, a statistically significant negative alpha

of -0.02% at a 10% significance level is noted for its integration into the value-weighted portfolio. This indicates that while variations in returns can be attributed to ESG considerations or other uncontrolled factors to some extent in different periods and for various portfolio allocation strategies, this relationship tends to approach zero when examining the entire period. When analysing the risk factors, a consistent trend emerged as sustainable portfolios showcased lower market factor exposure, pointing towards reduced systematic risk. Moreover, incorporating sustainability into equal-weighted portfolios resulted in reduced small-cap risk, whereas it exhibited a pronounced loading on the CMA factor for value-weighted portfolios, signifying a more conservative investment approach among its constituents.

Unexpectedly, introducing additional criteria into portfolios, such as ESGC weighting, resulted in higher annual turnover. The MKTV portfolio exhibited the lowest turnover rate at an average of 7.2%, followed by ESGV at 18.2%, MKTE at 24.0%, and ESGE at 28% (see Appendix 5). Table 10 displays the regression outcomes by integrating round-trip transaction costs within the spectrum of 50 to 200 basis points, which is in line with the methodology outlined by Kempf & Osthoff (2007). The heightened turnover rate of ESGV led to a reduction in risk-adjusted returns when compared to the value-weighted MKTV portfolio. However, deviating from an equal-weighted index by incorporating sustainability criteria no longer imposed a significant additional drag on risk-adjusted returns. This is attributed to both portfolios exhibiting increased turnover rates, thereby incurring higher transaction costs simultaneously. Consequently, when accounting for transaction costs, the ESGV portfolio yielded a negative monthly alpha ranging between -0.03% and -0.04% at significance levels between 5% and 1%, depending on the magnitude of the round-trip transaction cost.

**Table 10 :** Fama-French 5-factor loading on ESGE – MKTE and ESGV – MKTV difference portfolios, incorporating turnover costs, with S&P 500 constituents

	50 bp		100 bp		150 bp		200 bp	
	ESGE - MKTE	ESGV - MKTV	ESGE - MKTE	ESGV - MKTV	ESGE - MKTE	ESGV - MKTV	ESGE - MKTE	ESGV - MKTV
RMt-Rft	-0.0079*** (0.0028)	-0.0102** (0.0045)	-0.0079*** (0.0028)	-0.0104** (0.0045)	-0.0079*** (0.0028)	-0.0107** (0.0045)	-0.0080*** (0.0028)	-0.0109** (0.0046)
SMB	-0.0142*** (0.0049)	0.0091 (0.0068)	-0.0138*** (0.0049)	0.0101 (0.0070)	-0.0133*** (0.0049)	0.0112 (0.0073)	-0.0129*** (0.0049)	0.0122 (0.0077)
HML	0.0016 (0.0048)	0.0056 (0.0083)	0.0015 (0.0049)	0.0053 (0.0084)	0.0013 (0.0049)	0.0051 (0.0087)	0.0011 (0.0050)	0.0049 (0.0089)
RMW	0.0100* (0.0060)	0.0070 (0.0121)	0.0107* (0.0060)	0.0090 (0.0121)	0.0114* (0.0061)	0.0109 (0.0123)	0.0121* (0.0061)	0.0129 (0.0124)
CMA	0.0073 (0.0074)	0.0380*** (0.0141)	0.0072 (0.0074)	0.0379*** (0.0144)	0.0071 (0.0075)	0.0378** (0.0147)	0.0070 (0.0075)	0.0377** (0.0150)
Alpha	0.0001 (0.0001)	-0.0003** (0.0001)	0.0000 (0.0001)	-0.0003** (0.0001)	0.0000 (0.0001)	-0.0004*** (0.0001)	0.0000 (0.0001)	-0.0004*** (0.0001)
Observations	168	168	168	168	168	168	168	168
Adj. R <sup>2</sup>	0.2507	0.2312	0.2481	0.2288	0.2447	0.2248	0.2407	0.2194

Note: This table illustrates the Fama-French 5-factor loading on the difference portfolios ESGE – MKTE (formed by buying ESGE and selling MKTE) and ESGV – MKTV (formed by buying ESGV and selling MKTV) observed between 2010 and 2023. These portfolios are constructed using constituents from the S&P 500 as the underlying basis. The data is organised into four distinct levels of round-trip transaction costs, ranging from 50 basis points to 200 basis points. HAC robust standard errors are presented in parentheses. Levels of statistical significance are denoted by \*\*\* (1%), \*\* (5%), and \* (10%).

### 6.1.4 Robustness Tests (U.S.)

Subsidiary regressions were run to enhance the analysis's robustness. Initially, the results were tested for value-weighted top and bottom portfolios (T20V and B20V), followed by the alternative Carhart 4-factor regression model. Finally, all regressions were rerun with statistical ESGC score outliers excluded.

Appendix 2.1 details the regression tables for the market value-weighted T20V and B20V portfolios. Value-weighted portfolios showed more pronounced outcomes, which can be attributed to specific stocks exerting a substantial influence on returns. The inherent low diversification, with only 20% of the index represented in each portfolio, contributed to these outcomes. Contradictory findings emerged, with significantly negative alphas for T20V and some positive but non-significant alphas for B20V. This led to a significantly negative monthly alpha of -0.32% in a long-short strategy at a 5% confidence level during the entire observation period, primarily driven by B20V outperforming T20V from 2010 to 2016. This aligns with the observed negative alpha from investing in a long ESGV short MKTV portfolio during that period. Distinctions in factor loadings were evident as well. Results during the COVID-19 sub-period remained overall consistent, with a small positive alpha at a 10% confidence interval for sustainable portfolios during the drawdown period and no statistically significant alphas in the subsequent recovery period.

Appendix 2.2 summarises all preceding regressions against the Carhart 4-factor model. Shifts in alpha significance levels could be observed. Only two alphas maintained their significance. One is a positive alpha in the T20E B20E difference portfolio during the recent subperiod from 2017 to 2023. Similarly, a positive alpha persisted in the long ESGE short MKTE portfolio, once again in the recent subperiod. All other previously mentioned alphas were no longer statistically significant, including the negative alpha associated with integrating ESGC weighting in value-weighted portfolios, as returns were partially explained by the momentum factor. However, the Carhart 4-factor model didn't capture returns as well as the Fama-French 5-factor model, as evidenced by a lower adjusted R<sup>2</sup> ratio.

The final analysis repeated the regressions after excluding previously overlooked ESGC score outliers (refer to Appendix 7). As anticipated and supported by Appendix 2.3, the removal of only 2 statistical outliers across 2 respective periods did not substantially alter the results; there were no changes in significance levels, although a slight increase in alpha was noted in the T20E B20E difference portfolio in the subperiod from 2017 to 2023.

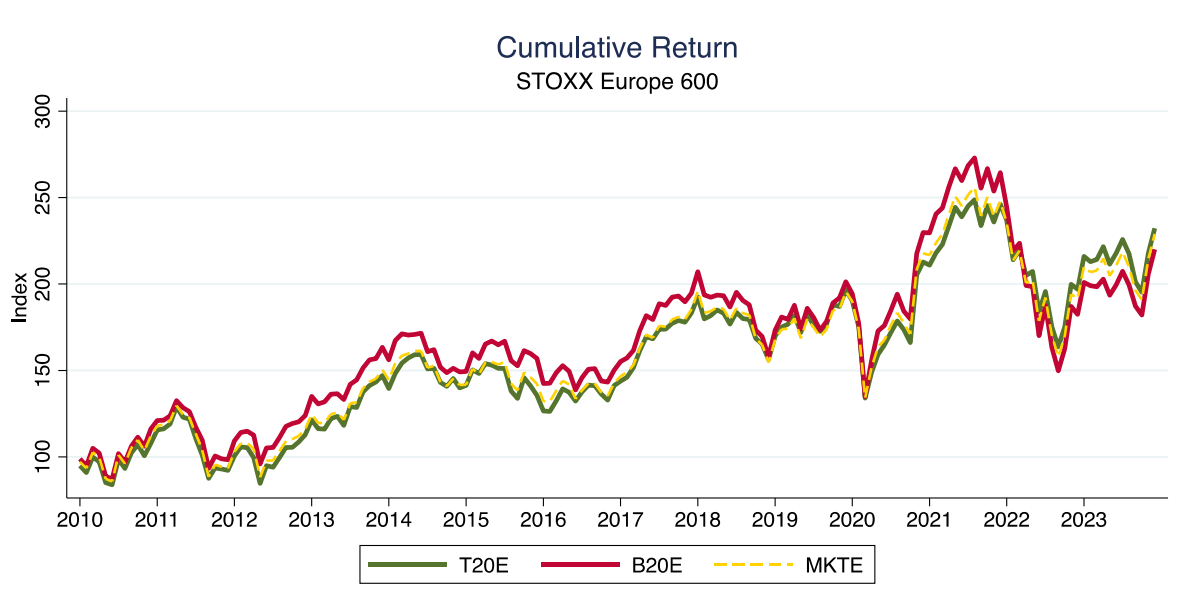
The robustness analysis indicated that results could significantly differ based on study methodology. The robustness of certain alpha findings across diverse portfolio strategies and regression methodologies was limited. In the broader context of the United States, overall market efficiency was apparent, with only a few positive alphas identified in the recent subperiod for equal-weighted and simple ESGC-weighted portfolios, respectively. However, the conflicting results from value-weighted portfolios introduced complexity, leading to mixed overall conclusions, which are further discussed in Chapter 7.

## 6.2 Europe

### 6.2.1 Top/Bottom Portfolio Analysis (Europe)

In the subsequent phase, the European market is analysed through portfolios constructed based on STOXX Europe 600 index constituents. Table 11 presents the descriptive performance statistics of the equal-weighted T20E and B20E portfolios, with MKTE serving as the benchmark. Comparable to the U.S., the more sustainable portfolio outperformed over the entire observation period, with annualised returns of 6.20% compared to 5.80% for B20E, alongside lower volatility and a higher Sharpe Ratio of 25.5% against 22.5% for B20E. Additionally, as illustrated in Figure 6, a significant return divergence emerged in the post-2021 period, with the T20E portfolio yielding considerably higher returns.

**Figure 6 :** Cumulative Returns for equal-weighted T20E, B20E and MKTE portfolios based on STOXX Europe 600 constituents



Compared to the U.S. scenario, more pronounced shifts in risk-adjusted returns across subperiods were observed. Between 2010 and 2016, the B20E portfolio demonstrated superior risk-adjusted returns, while from 2017 to 2023, it was the T20E portfolio.

In the earlier period, B20E achieved an annualised return of 5.27%, surpassing T20E's 4.14%. B20E accomplished this with lower volatility at 19.44%, compared to T20E's 20.30%. This finding contradicts the typical trend observed in all other results of this study, where portfolios with higher sustainability criteria tend to experience lower volatility. Consequently, the B20E portfolio attained a Sharpe ratio of 27.10%, surpassing T20E by a significant margin, which achieved a Sharpe ratio of only 20.4%.

During the more recent period from 2017 to 2023, the T20E portfolio outperformed with significantly higher annualised returns of 7.36% compared to B20E's 5.62%. It also maintained lower volatility, leading to a substantially higher Sharpe ratio of 26% for T20E, in contrast to only 16.1% for B20E.

**Table 11 :** Performance statistic measures for equal-weighted T20E, B20E and MKTE portfolios based on STOXX Europe 600 constituents

	2010-2023			2010-2016			2017-2023		
	T20E	B20E	MKTE	T20E	B20E	MKTE	T20E	B20E	MKTE
Annualised Return	6.20%	5.80%	6.10%	4.14%	5.27%	4.42%	7.36%	5.62%	6.98%
Annualised Volatility	20.92%	21.81%	20.89%	20.30%	19.44%	19.69%	21.63%	24.07%	22.13%
Sharpe Ratio	25.5%	22.5%	25.0%	20.4%	27.1%	22.4%	26.0%	16.1%	23.7%

Interestingly, when comparing the T20E and B20E portfolios to the market index MKTE, it suggests that the T20E portfolio didn't necessarily underperform the market in the subperiod 2010-2016 or outperform the market in the subperiod 2017-2023. Instead, it appears that the B20E portfolio outperformed and underperformed in those subperiods.

**Table 12 :** Fama-French 5-factor loading on equal-weighted T20E, B20E and L-S portfolios based on STOXX Europe 600 constituents

	2010 - 2023			2010 - 2016			2017 - 2023		
	T20E	B20E	L-S	T20E	B20E	L-S	T20E	B20E	L-S
RMt-Rft	1.1029*** (0.0187)	1.1350*** (0.0221)	-0.0321 (0.0232)	1.0725*** (0.0337)	1.0783*** (0.0315)	-0.0059 (0.0373)	1.0955*** (0.0242)	1.1725*** (0.0263)	-0.0770** (0.0297)
SMB	0.0457 (0.0459)	0.4707*** (0.0445)	-0.4250*** (0.0538)	-0.0608 (0.0501)	0.3996*** (0.0586)	-0.4604*** (0.0653)	0.1611** (0.0652)	0.4109*** (0.0734)	-0.2498*** (0.0734)
HML	0.2215*** (0.0539)	0.1007* (0.0602)	0.1208 (0.0734)	0.2396*** (0.0670)	-0.0343 (0.1376)	0.2739** (0.1297)	0.2321*** (0.0748)	0.1603** (0.0765)	0.0717 (0.1037)
RMW	0.1801*** (0.0652)	0.1705* (0.0942)	0.0096 (0.0998)	0.1291 (0.0856)	-0.0984 (0.1590)	0.2275 (0.1597)	0.2073** (0.0974)	0.2213** (0.1107)	-0.0140 (0.1322)
CMA	-0.0456 (0.0785)	-0.2575** (0.1031)	0.2120* (0.1193)	-0.0040 (0.1253)	-0.2647* (0.1420)	0.2607* (0.1542)	-0.0823 (0.1100)	-0.3039** (0.1492)	0.2216 (0.1801)
Alpha	-0.0006 (0.0006)	-0.0013 (0.0009)	0.0007 (0.0008)	0.0008 (0.0009)	0.0011 (0.0009)	-0.0002 (0.0013)	-0.0013* (0.0008)	-0.0028** (0.0012)	0.0015 (0.0011)
Observations	168	168	168	84	84	84	84	84	84
Adj. R <sup>2</sup>	0.9789	0.9665	0.3831	0.9776	0.9643	0.4454	0.9815	0.9703	0.3609

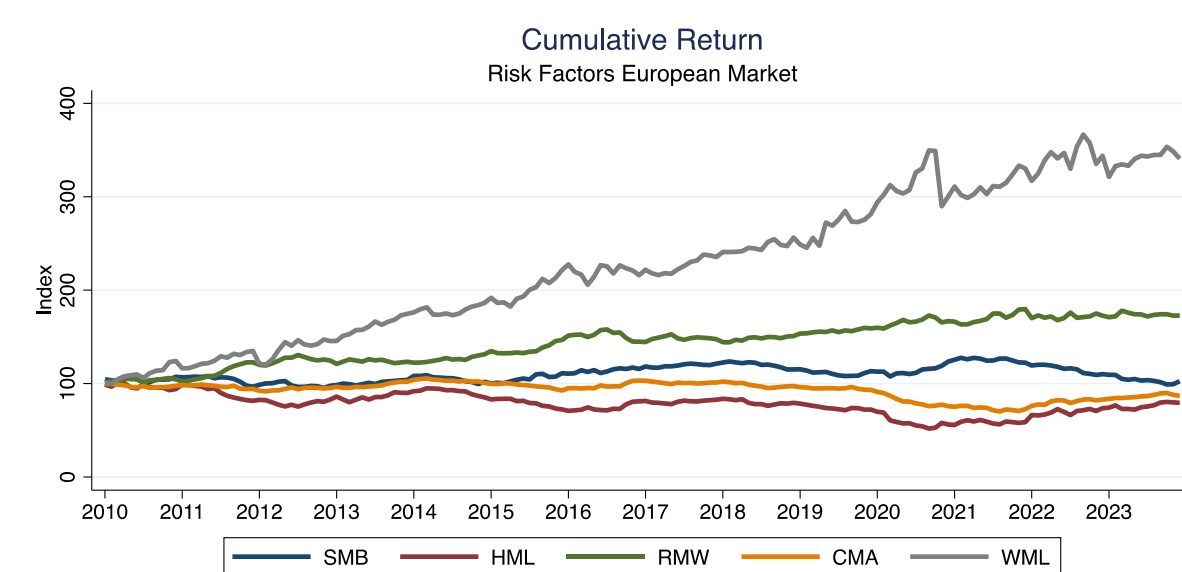
Note: This table illustrates the Fama-French 5-factor loading on the equal-weighted T20E, B20E, and L-S Portfolios, constructed using constituents from the STOXX Europe 600 as the underlying basis. The data is categorised into three distinct time periods: 2010-2023, 2010-2016, and 2017-2023. HAC robust standard errors are presented in parentheses. Levels of statistical significance are denoted by \*\*\* (1%), \*\* (5%), and \* (10%).

Table 12 displays the Fama-French 5-factor regression results for the T20E, B20E, and the long-short portfolio L-S across three distinct periods. Similar to the U.S., significant alphas were only observed from 2017 to 2023. During this timeframe, B20E exhibited a significantly negative monthly alpha of -0.28% at a 5% confidence interval. However, the T20E portfolio also recorded a negative alpha of -0.13% at a 10% confidence interval. Implementing a long-short strategy resulted in a positive abnormal return of 0.15% in the European context, but it lacked statistical significance. In the earlier period from 2010 to 2016, both T20E and B20E exhibited positive abnormal returns, with negligible differences, suggesting no clear over- or underperformance between the two portfolios.

In contrast to the U.S., the European portfolio displayed significant loadings on various risk factors, suggesting a stronger inclination of portfolios towards specific asset characteristics. Particularly, the European B20E portfolio demonstrated heightened exposure to systematic and small-cap risks compared to its U.S. counterpart. Additionally,

the T20E portfolio in Europe exhibited a greater inclination towards value stocks and adhered to a more conservative investment strategy relative to B20E.

**Figure 7 :** Cumulative Returns for Fama-French & Carhart risk factors in the European market



The increased differences in style allocation coincided with heightened consistency in style factor returns, explaining return differences to a greater extent than in the U.S. Figure 7 illustrates the cumulative returns of various risk factors, suggesting that, in comparison to the U.S. scenario, factor investing in Europe appeared more financially promising thanks to consistent trends. Investing in past winners and, to a lesser extent, companies with robust operational profitability would have yielded substantial profits. Leading up to the COVID-19 pandemic, favourable investment opportunities would have also included small-cap growth stocks and companies with aggressive investment strategies. Subsequently, the post-pandemic period favoured investments in large-cap value stocks with conservative investment strategies.

Although the descriptive statistics highlighted performance disparities, most of these variations can be attributed to various risk factors, primarily size, value, and investment. The initial outperformance of the B20E portfolio can be fully attributed to its higher allocation to small-cap aggressive growth stocks, which performed exceptionally well during that period. In summary, while no statistically significant differences are apparent, there is a clear indication that the recent time period favoured the sustainable portfolio.

### 6.2.2 COVID-19 Period (Europe)

When the COVID-19 pandemic period was examined, differences in comparison to the U.S. context became apparent, which manifested in both the descriptive statistics and the multi-factor regressions.

Table 13 outlines the descriptive performance statistics. In the recovery phase, the sustainable portfolio displayed lower risk-adjusted performance, as measured by the Sharpe and Sortino Ratio, mirroring the U.S. context. However, during the drawdown period, no significant discrepancies in downside risk and risk-adjusted performance between the top and bottom portfolios were observed. This contrasts with the U.S., where the sustainable portfolio yielded higher risk-adjusted returns during the drawdown.

**Table 13 :** Performance statistic measures for equal-weighted T20E, B20E and MKTE portfolios based on STOXX Europe 600 constituents during the COVID-19 pandemic

	Entire period			Drawdown period			Recovery period		
	T20E	B20E	MKTE	T20E	B20E	MKTE	T20E	B20E	MKTE
Average return	-0,05%	-0,01%	-0,02%	-0,91%	-0,91%	-0,87%	0,23%	0,29%	0,25%
Average volatility	2,62%	2,68%	2,51%	4,22%	4,35%	4,08%	1,77%	1,79%	1,67%
Average downside volatility	1,99%	2,01%	1,91%	3,50%	3,53%	3,37%	1,13%	1,13%	1,07%
Highest daily loss	-13,96%	-13,48%	-13,36%	-13,96%	-13,48%	-13,36%	-4,81%	-5,16%	-4,63%
Maximum drawdown	-39,74%	-41,21%	-38,96%	-39,74%	-41,21%	-38,96%	-8,53%	-4,90%	-4,89%
Sharpe Ratio	-1,89%	-0,31%	-1,00%	-21,66%	-21,18%	-21,52%	13,28%	16,00%	15,08%
Sortino Ratio	-2,37%	-0,31%	-1,21%	-25,88%	-25,86%	-25,80%	20,82%	25,45%	23,58%

Note: The term "Drawdown period" denotes the timeframe from 20/02/2020 to 07/04/2020, while the "Recovery period" spans from 08/04/2020 to 31/08/2020. The "Entire period" encompasses both the Drawdown and Recovery periods.

**Table 14 :** Fama-French 5-factor loading on equal-weighted T20E, B20E and L-S portfolios based on STOXX Europe 600 constituents during the COVID-19 pandemic

	Entire period			Drawdown period			Recovery period		
	T20E	B20E	L-S	T20E	B20E	L-S	T20E	B20E	L-S
RMt-Rft	1.1269*** (0.0322)	1.1836*** (0.0370)	-0.0567** (0.0256)	1.1857*** (0.0449)	1.2093*** (0.0731)	-0.0236 (0.0468)	1.1608*** (0.0695)	1.2588*** (0.0685)	-0.0981*** (0.0344)
SMB	0.1956** (0.0788)	0.4168*** (0.0735)	-0.2212*** (0.0626)	0.2555*** (0.0853)	0.4431*** (0.1278)	-0.1876* (0.1038)	0.2928** (0.1470)	0.5652*** (0.1756)	-0.2724*** (0.0979)
HML	0.2463** (0.1054)	0.2697* (0.1583)	-0.0234 (0.1285)	-0.0513 (0.1598)	0.0906 (0.3003)	-0.1419 (0.2310)	0.3715*** (0.1245)	0.3543** (0.1493)	0.0173 (0.1096)
RMW	0.1283 (0.1533)	-0.0193 (0.2780)	0.1476 (0.2953)	-0.0647 (0.2907)	-0.4705 (0.5989)	0.4057 (0.6749)	0.1870 (0.1865)	0.1791 (0.2449)	0.0079 (0.2094)
CMA	0.2123 (0.1688)	0.0640 (0.2185)	0.1483 (0.1542)	0.1329 (0.2128)	-0.0477 (0.2798)	0.1806 (0.1964)	0.1197 (0.2043)	0.0298 (0.2724)	0.0899 (0.1963)
Alpha	0.0004 (0.0004)	0.0008* (0.0005)	-0.0004 (0.0004)	0.0006 (0.0012)	0.0017 (0.0015)	-0.0011 (0.0010)	0.0001 (0.0004)	0.0001 (0.0004)	-0.0000 (0.0004)
Observations	138	138	138	34	34	34	104	104	104
Adj. R <sup>2</sup>	0.9620	0.9545	0.1091	0.9780	0.9672	0.1269	0.9311	0.9252	0.0672

Note: This table illustrates the Fama-French 5-factor loading on the equal-weighted T20E, B20E, and L-S Portfolios, constructed using constituents from the STOXX Europe 600 as the underlying basis. The data is categorised into three distinct time periods: "Entire period" 20/02/2020-31/08/2020, "Drawdown period" 20/02/2020 – 07/04/2020, and "Recovery period" 08/04/2020 – 31/08/2020. HAC robust standard errors are presented in parentheses. Levels of statistical significance are denoted by \*\*\* (1%), \*\* (5%), and \* (10%).

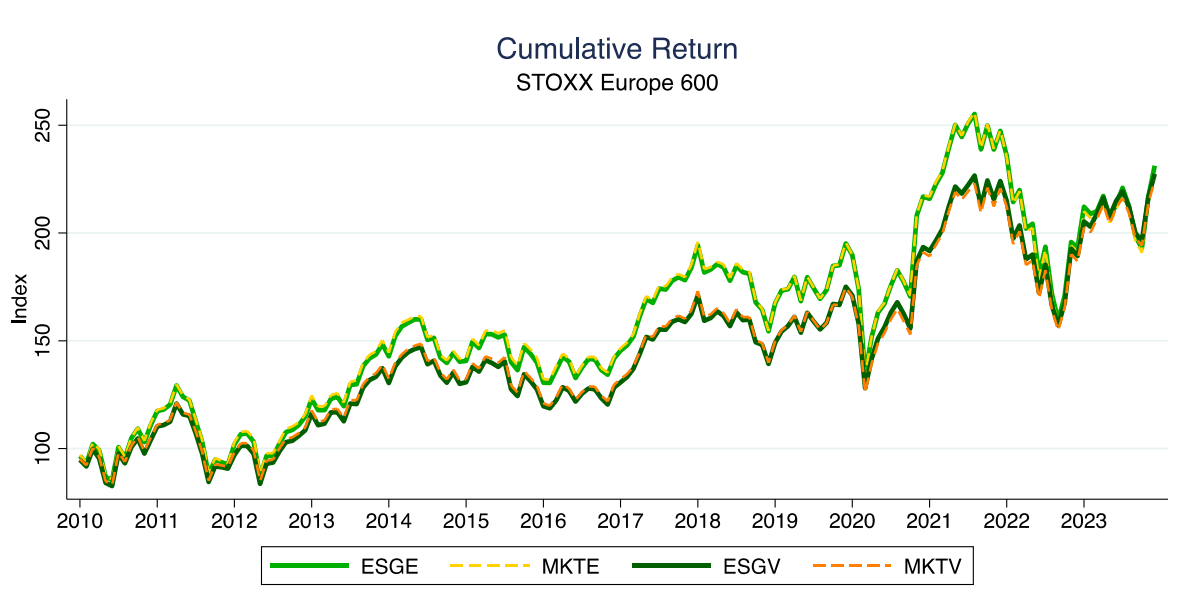
Table 14 presents the Fama-French regression results for the COVID-19 pandemic period. In contrast to the U.S., the drawdown period showed a higher daily alpha of 0.17% for B20E, although not statistically significant, compared to 0.06% for T20E. This suggests that, during the drawdown, the low sustainability portfolio outperformed the high sustainability index by a cumulative 3.7%, although the lack of statistical significance raises the possibility of chance. This finding contradicts the U.S., where the high

sustainability portfolio performed better during the drawdown. During the recovery period, there were no discernible differences in abnormal returns between the Top and Bottom portfolios, mirroring the U.S. scenario. This implies that the variations in return can be ascribed to the Fama-French risk factors, particularly the SMB factor. Over the entire period, a statistically significant daily alpha of 0.08% at a 10% confidence interval was noted for B20E, but the distinction with the T20E portfolio was no longer significant. Thus, the European context suggests that sustainable portfolios may underperform during drawdowns but not during recovery periods.

### 6.2.3 ESG-weighted Portfolio Analysis (Europe)

In the final analysis, this study examined the four portfolios based on different asset allocation strategies. Figure 8 depicts the cumulative returns. The returns of the ESGC-weighted portfolio (ESGE) closely mirrored those of the equal-weighted market portfolio (MKTE), and those of the ESGC and value-weighted portfolio (ESGV) mirrored those of the value-weighted market portfolio (MKTV).

**Figure 8 :** Cumulative Returns for ESGE, MKTE, ESGV and MKTV portfolios based on STOXX Europe 600 constituents



This alignment is evident in both return and volatility patterns, as indicated in Table 15, resembling the situation observed in the United States. However, in the European context, this consistency was notably more pronounced. Once again, the value-weighted portfolios exhibited lower volatility than their equal-weighted counterparts. When comparing the sustainable portfolios to their respective benchmarks, no significant differences were evident. Nevertheless, both ESGE and ESGV portfolios demonstrated slightly higher returns and marginally lower volatility, resulting in slightly higher Sharpe ratios than their benchmarks.

**Table 15 :** Performance statistic measures for ESGE, MKTE, ESGV and MKTV portfolios based on STOXX Europe 600 constituents

	2010-2023				2010-2016				2017-2023			
	ESGE	MKTE	ESGV	MKTV	ESGE	MKTE	ESGV	MKTV	ESGE	MKTE	ESGV	MKTV
Annualised Return	6.2%	6.1%	6.0%	5.9%	4.3%	4.4%	2.7%	2.9%	7.2%	7.0%	8.5%	8.1%
Annualised Volatility	20.9%	20.9%	19.3%	19.4%	19.8%	19.7%	19.1%	18.9%	22.0%	22.1%	19.7%	19.9%
Sharpe Ratio	25.4%	25.0%	26.8%	26.2%	21.7%	22.4%	14.1%	15.2%	25.0%	23.7%	34.5%	32.2%

Akin to the U.S. scenario, incorporating ESGC weights into the portfolio significantly improved sustainability characteristics. This is demonstrated by an additional increase of 5.1 ESGC points for the ESGE portfolio compared to MKTE and 4.1 ESGC points for the ESGV portfolio compared to the MKTV (see Appendix 4).

Table 16 displays the Fama-French five-factor regression results for a long ESGE short MKTE Portfolio and a long ESGV short MKTV portfolio across different time periods. Significant monthly alphas were only evident in the recent period, with the ESGV portfolio surpassing the MKTV portfolio by 0.04% at a 1% confidence level and the ESGE portfolio outperforming the MKTE portfolio by 0.02% at a 10% confidence interval. However, over the entire period, no statistically significant alphas were observed, though the more sustainable portfolios exhibited a positive alpha of 0.01% in this sample. These results demonstrate a greater consistency in the positive relationship compared to the mixed findings in the U.S. context.

**Table 16 :** Fama-French 5-factor loading on ESGE – MKTE and ESGV – MKTV difference portfolios using STOXX Europe 600 constituents

	2010 - 2023		2010 - 2016		2017 - 2023	
	ESGE - MKTE	ESGV - MKTV	ESGE - MKTE	ESGV - MKTV	ESGE - MKTE	ESGV - MKTV
RMt-Rft	-0.0002 (0.0020)	0.0008 (0.0030)	0.0014 (0.0039)	0.0039 (0.0043)	-0.0052** (0.0021)	-0.0078*** (0.0027)
SMB	-0.0468*** (0.0055)	-0.0350*** (0.0068)	-0.0537*** (0.0074)	-0.0302*** (0.0068)	-0.0257*** (0.0065)	-0.0163* (0.0083)
HML	0.0116 (0.0072)	-0.0290*** (0.0109)	0.0312* (0.0170)	0.0268 (0.0168)	0.0076 (0.0080)	-0.0412*** (0.0080)
RMW	0.0066 (0.0102)	-0.0244* (0.0145)	0.0349 (0.0218)	0.0368 (0.0223)	0.0000 (0.0112)	-0.0234 (0.0142)
CMA	0.0222* (0.0126)	0.0147 (0.0138)	0.0368** (0.0169)	0.0371** (0.0144)	0.0143 (0.0152)	0.0063 (0.0130)
Alpha	0.0001 (0.0001)	0.0001 (0.0001)	-0.0001 (0.0002)	-0.0002 (0.0002)	0.0002* (0.0001)	0.0004*** (0.0001)
Observations	168	168	84	84	84	84
Adj. R <sup>2</sup>	0.3802	0.1779	0.4685	0.3110	0.3399	0.4822

Note: This table illustrates the Fama-French 5-factor loading on the difference portfolios ESGE – MKTE (formed by buying ESGE and selling MKTE) and ESGV – MKTV (formed by buying ESGV and selling MKTV). These portfolios are constructed using constituents from the STOXX Europe 600 as the underlying basis. The data is categorised into three distinct time periods: 2010-2023, 2010-2016, and 2017-2023. HAC robust standard errors are presented in parentheses. Levels of statistical significance are denoted by \*\*\* (1%), \*\* (5%), and \* (10%).

Examining the factor loadings revealed that during the period from 2017 to 2023, the sustainable portfolios exhibited lower exposure to systematic risk, evidenced by a significant negative market factor. Moreover, all sustainable portfolios leaned towards

larger companies. ESGE preferred value stocks compared to MKTE, while ESGV demonstrated a tendency toward growth stocks compared to MKTV. Furthermore, all sustainable portfolios adopted a more conservative investment strategy, as indicated by a positive CMA factor loading.

Concerning annual turnover, the MKTV portfolio boasted the lowest rate at an average of 8%, followed by ESGV at 18.8%, MKTE at 30.08%, and ESGE at 32.3% (see Appendix 5). Table 17 presents the regression results upon introducing round-trip transaction costs within the range of 50 to 200 basis points, following Kempf & Osthoff's (2007) approach.

**Table 17 :** Fama-French 5-factor loading on ESGE – MKTE and ESGV – MKTV difference portfolios, incorporating turnover costs, with STOXX Europe 600 constituents

	50 bp		100 bp		150 bp		200 bp	
	ESGE - MKTE	ESGV - MKTV	ESGE - MKTE	ESGV - MKTV	ESGE - MKTE	ESGV - MKTV	ESGE - MKTE	ESGV - MKTV
RMt-Rft	-0.0002 (0.0020)	0.0006 (0.0031)	-0.0003 (0.0020)	0.0003 (0.0032)	-0.0003 (0.0020)	0.0001 (0.0033)	-0.0004 (0.0020)	-0.0002 (0.0034)
SMB	-0.0470*** (0.0054)	-0.0356*** (0.0067)	-0.0471*** (0.0054)	-0.0361*** (0.0067)	-0.0472*** (0.0054)	-0.0367*** (0.0067)	-0.0473*** (0.0054)	-0.0373*** (0.0069)
HML	0.0118 (0.0072)	-0.0283** (0.0110)	0.0120* (0.0072)	-0.0275** (0.0111)	0.0122* (0.0072)	-0.0268** (0.0113)	0.0124* (0.0072)	-0.0260** (0.0116)
RMW	0.0068 (0.0102)	-0.0221 (0.0144)	0.0070 (0.0102)	-0.0199 (0.0145)	0.0072 (0.0102)	-0.0176 (0.0147)	0.0074 (0.0102)	-0.0154 (0.0150)
CMA	0.0217* (0.0125)	0.0138 (0.0138)	0.0212* (0.0125)	0.0128 (0.0139)	0.0208* (0.0125)	0.0119 (0.0142)	0.0203 (0.0124)	0.0110 (0.0147)
Alpha	0.0001 (0.0001)	0.0001 (0.0001)	0.0001 (0.0001)	0.0000 (0.0001)	0.0001 (0.0001)	0.0000 (0.0001)	0.0000 (0.0001)	-0.0000 (0.0001)
Observations	168	168	168	168	168	168	168	168
Adj. R <sup>2</sup>	0.3801	0.1821	0.3799	0.1838	0.3795	0.1833	0.3790	0.1808

Note: This table illustrates the Fama-French 5-factor loading on the difference portfolios ESGE – MKTE (formed by buying ESGE and selling MKTE) and ESGV – MKTV (formed by buying ESGV and selling MKTV) observed between 2010 and 2023. These portfolios are constructed using constituents from the STOXX Europe 600 as the underlying basis. The data is organised into four distinct levels of round-trip transaction costs, ranging from 50 basis points to 200 basis points. HAC robust standard errors are presented in parentheses. Levels of statistical significance are denoted by \*\*\* (1%), \*\* (5%), and \* (10%).

As expected, the alpha decreased, but no negative or statistically significant alphas emerged. Once again, ESGV's higher turnover compared to MKTV influenced risk-adjusted returns to the greatest extent. No significant differences were evident in the European context, even after considering transaction costs.

#### 6.2.4 Robustness Tests (Europe)

Comparable to the U.S. analysis, robustness checks were conducted to address over-reliance on a specific portfolio strategy or multi-factor regression.

Appendix 3.1 details the regression tables for value-weighted top and bottom portfolios labelled T20V and B20V. Once again, value-weighted portfolios exhibited significantly more pronounced outcomes, through certain companies exerting a considerable influence on portfolio returns and contributing to lower diversification. In the European context,

the robustness analysis remained consistent, revealing an exceptionally large positive alpha of 0.43% at a 1% confidence level during the recent subperiod when employing a long-short strategy. No notable differences were observed in other subperiods or the overall observation period. Additionally, results remained relatively consistent during the COVID-19 subperiod, revealing that low sustainability portfolios outperformed through a significant negative alpha of -0.07% when investing in a long-short portfolio throughout the entire observation period at a 10% confidence interval.

Appendix 3.2 presents the regression tables employing the Carhart 4-factor model. Previously significant alphas lost their significance due to statistically significant loadings on the momentum factor. Despite this alteration, the overall trend persisted, although without statistical significance.

In the final analysis, companies identified as statistical outliers due to excessively low ESGC scores were omitted (see Appendix 7). As anticipated and supported by Appendix 3.3, removing these outliers did not substantially alter the results, although minor adjustments in alpha levels were discernible. Notably, the B20E portfolio was the most influenced by this adjustment, with its monthly alpha experiencing a marginal increase of 0.02% over the entire period. However, no changes in significance levels were observed in any of the portfolio tests.

In contrast to the U.S. context, this robustness analysis didn't highlight variations in findings but rather emphasised changes in significance levels based on the employed methodology. The consistent trend is that sustainable portfolios exhibited a modest positive alpha during the recent subperiod but a negative alpha during the COVID-19 drawdown period. However, the statistical significance of these observations varied depending on the regression model used.

## 7 Discussion

This chapter presents the discussion and comparison of the empirical findings with the existing literature. The primary objective is to address the central research question of this study: whether the integration of ESG criteria into equity portfolios affects risk-adjusted returns and, consequently, how investors might formulate strategies to capitalise on this information. This paper takes a reflective stance, revisits the findings, theories and literature mentioned earlier, and quantifies the implications for an investor.

### 7.1 ESG and Market Efficiency

Existing literature posits the overall efficiency of markets, asserting that publicly available information is promptly incorporated into asset prices. Nevertheless, some studies suggest the presence of anomalies, raising the question of whether ESG information is entirely reflected in market prices or if abnormal returns can be earned through its utilisation. This discussion touches upon the domain of behavioural finance, examining the motivations and preferences of individual investors.

In the initial analysis, focusing on extreme scenarios where investors solely allocate funds to assets with the highest or lowest sustainability scores, significant differences in returns and volatility emerged. The sustainable portfolio generally exhibited lower volatility and delivered higher raw returns in the recent subperiod. Consequently, investors would have accrued greater cumulative returns by investing in sustainable portfolios with reduced associated volatility risk. Some of these return variations could be attributed to diverse portfolio characteristics beyond ESG considerations and the risk factors to which they are exposed to.

Consistent with previous research (Humphrey et al., 2012; Halbritter & Dorfleitner, 2015; Teti et al., 2023; Bannier et al., 2019), portfolios with high sustainability scores demonstrated reduced exposure to market risk and were considerably larger, particularly pronounced in the European context compared to the U.S. These portfolios also exhibited a higher allocation to value stocks (Humphrey et al., 2012; Halbritter & Dorfleitner, 2015) and tended to follow a more conservative investment strategy in Europe. Therefore, differences in returns could be attributed to variations in factor returns, with no indication of a positive alpha over the entire period and limited evidence of a positive alpha in the recent subperiod. This alpha remained robust across various multi-factor regressions in the U.S. yet turned negative for value-weighted indexes. Conversely, this positive alpha remained robust for value-weighted indexes in Europe but lacked significance in different multi-factor regressions.

This paper found that there were no statistically significant differences in risk-adjusted returns between the two portfolios throughout the entire observation period. This implies that, overall, the market effectively incorporates sustainability characteristics and their potential impact on future company performance. This conclusion is in line with findings from studies such as Bello (2005), Humphrey et al. (2012), Halbritter & Dorfleitner (2015),

and Milonas et al. (2022), suggesting that sustainability characteristics only result in different factor exposures, rather than leading to distinct risk-adjusted returns, particularly when considering long-term investments. Consequently, these findings support the market efficiency hypothesis.

This study contradicts previous research asserting consistently positive (Kempf & Osthoff, 2007; Statman & Glushkov, 2009; Steen et al., 2020; Abate et al., 2021) or negative (Bauer et al., 2005; Bannier et al., 2019; Das et al., 2018) alphas through the examination of sustainable and unsustainable portfolios. The discrepancies with this study could be attributed to differences in methodology. A considerable portion of these studies predominantly examined mutual fund data (Abate et al., 2021; Bauer et al., 2005; Das et al., 2018; Steen et al., 2020). This approach introduces inherent limitations, as discussed in Chapter 5.1, and adds complexity due to variations in the methodology employed to assess risk-adjusted returns. Abate et al. (2021) utilised DEA efficiency models, Steen et al. (2020) applied the Fama-French 3-factor model, potentially overlooking other significant risk factors, Das et al. (2018) used the Sharpe ratio while controlling for different fund characteristics such as management tenure, expense ratio, age, and fund size, and finally, Bauer et al. (2005) adopted the Carhart 4-factor model. This diversity in methodologies underscores the ongoing challenge researchers face in determining which models best capture the risks associated with different mutual funds, contributing to the difficulty in establishing consistent findings. Other studies, such as those conducted by Kempf & Osthoff (2007) and Statman & Glushkov (2009), are already dated and had different research objectives. They examined various screening strategies on portfolio returns and found positive alpha only for the best-in-class screening method.

Of particular interest, Bannier et al. (2019) employed a comparable methodology by constructing portfolios from constituents in both the American and European markets. They utilised data from Datastream, incorporating the Refinitiv ESG rating, and analysed these portfolios using the Carhart 4-factor regression. Their findings revealed an excess return for the low sustainability portfolio. This paper argues that these divergent findings may arise from several disparities, including the consideration of an earlier period and the utilisation of a different sample to represent the American and European markets. However, the primary distinction lies in their use of Refinitiv ESG ratings rather than ESGC ratings. It is advisable to prioritise the ESGC score, as it provides “a comprehensive evaluation of the company’s sustainability impact and conduct over time” (Refinitiv, 2022). Considering the consistent findings regarding the impact of sustainability on corporate financial performance as outlined in Chapter 4.1 and assuming its integration into market pricing in accordance with efficient markets, choosing an ESG score without controversies could lead to disregarding essential information with potential market implications. Additionally, the ESGC score enjoyed broad adoption in recent academic literature, as evidenced in studies such as Petridis et al. (2023) and Teti et al. (2023). This emphasises the mounting challenge in the ESG landscape, where stakeholders and researchers grapple with many divergent sustainability ratings, as discussed in Chapter 2.2.

Remarkably, the empirical analysis highlighted some deviations from market efficiency in the recent subperiod, aligning with recent studies by Teti et al. (2023) and De Spiegeleer et al. (2023). In both geographical regions, the findings revealed either statistically significant positive alphas for sustainable portfolios or long-short portfolios, or negative alphas for unsustainable portfolios (except for the value-weighted portfolios in the U.S. market). While the statistical robustness of these results varied across different regression models and portfolio allocation strategies, it is possible that the incorporation of sustainability criteria led to outperformance in the recent subperiod, contradicting the efficient market hypothesis. Potential explanations may be rooted in behavioural finance theories or limitations within this study, which are discussed in Chapter 7.4.

This alpha in the recent subperiod may be attributed to significant shifts in the regulatory landscape, heightened concerns about sustainability, a stock market crash, and increased media attention in recent years. These factors might have prompted changes in investor preferences or altered their overall perception of the importance of sustainability criteria, potentially influencing shifts in behaviour. Following Pedersen et al.'s (2021) model, this shift might indicate a transition from ESG-unaware to ESG-aware or even ESG-motivated investors, or as proposed by Hvidkjær (2017) a diminished underreaction to ESG information over time. This could have led to a continued preference for sustainable stocks, or an avoidance of sin stocks or those with high controversies, leading to changes in market prices. While behavioural factors might explain the presence of a positive alpha in the recent subperiod, its long-term sustainability appears doubtful.

Suppose there has been a shift from ESG-unaware to ESG-aware investors. In such a scenario, sustainable stocks were likely undervalued previously, as the positive impacts of ESG on their future cash flows were not fully reflected in their prices. Consequently, if investors and society at large become more ESG-aware, they would recognise this undervaluation and continuously buy these stocks, leading to increased demand and driving up their prices until they reach fair value. This continuous price appreciation during the transition period would have resulted in higher capital gains for investors, potentially explaining a positive alpha. If this scenario holds true, the transition period might have already ended, and once asset prices align with their fair value, further outperformance becomes highly improbable.

Alternatively, consider a scenario with a shift from ESG-aware to ESG-motivated investors. In this case, sustainable stocks were already correctly priced, but a surge in demand occurred not because of mispricing but due to societal preferences shifting towards sustainability. As societal awareness of social and environmental issues grows, along with a recognition of the imperative to contribute actively, a distinct trend or hype around sustainability may have emerged. This trend would be evident not only among individual investors but also at the institutional level, possibly influenced by evolving financial regulations. The underlying principle lies in the concept that investors perceive additional value beyond financial returns when investing in sustainable assets. This heightened demand for sustainable stocks would have driven up their prices, thereby generating a positive alpha in the recent period through capital gains. However, negative

alphas are anticipated in subsequent periods, as investor returns encompass both capital gains and distributions. An overpriced asset may fail to generate adequate distributions to justify its valuation, potentially leading to diminished future performance and ultimately realigning with market efficiency over the long term.

In summary, opting for integrating sustainability criteria into investment decisions would not have adversely affected investors; on the contrary, it could have marginally improved their risk-adjusted returns. Therefore, achieving both financial success and a positive impact is feasible, provided there are practical and cost-effective strategies for investors to integrate sustainability characteristics into their portfolios. Investors should perceive sustainability criteria not as a pathway to secure abnormal returns - given the unlikely persistence of the recently identified positive alpha - but rather as a strategy to allocate capital to more sustainable companies without sacrificing returns. This approach enhances overall utility without compromising financial gains. Overall, one can consider markets efficient, and integrating sustainability would not negatively affect an investor.

## **7.2 Exploration of the COVID-19 Period**

As this study highlights, investors neither reap continued rewards nor face penalties for holding sustainable assets over extended periods. Several researchers have suggested that sustainable assets offer asymmetric benefits, providing decreased downside risk during financial turmoil. However, this comes at the cost of an "insurance premium" (Lins et al., 2017; Engle et al., 2020; Becchetti et al., 2015; Bannier et al., 2019; Albuquerque et al., 2020). Based on this statement, sustainable assets are perceived as less risky, leading individuals to accept a premium through lower returns in normal periods. As outlined in Chapter 7.1, no evidence of such an insurance premium can be found, as overall risk-adjusted returns are not significantly different. The remaining question pertains to whether these assets can provide an extra layer of downside risk protection or increased risk-adjusted returns, especially in the context of the stock market crash triggered by the COVID-19 pandemic.

Becchetti et al. (2015) and Broadstock et al. (2021) suggest that enhanced risk-adjusted returns are particularly evident during the drawdown period, not in other periods. This study aligns with these suggestions to some extent. Specifically, there were no discernible differences in risk-adjusted returns during the recovery period, whereas differences were observed during the drawdown period, albeit without statistical significance. Thus, this paper supports the notion that divergent risk-adjusted returns become apparent amid heightened volatility, indicating that investors may experience varied outcomes mainly in times of increased market uncertainty. Evidence for enhanced risk-adjusted returns during the drawdown period was inconclusive, with variations observed across geographical regions and without statistically significant findings. While U.S. sustainable portfolios exhibited positive alphas in the drawdown, European counterparts showed negative alphas, making it challenging to definitively assess the impact of ESG during crises. Discrepancies may arise from ESG rating limitations, the specific context of the COVID-19 crisis, or the relatively short observation period. However, a study by Bannier

et al. (2019) indicated a positive relationship during crisis periods in the U.S. and a neutral one in Europe. This suggests that the perceived stabilising impact of ESG during times of high market volatility may be a phenomenon unique to the U.S., with varying perceptions among investors in different geographical regions. Additionally, the European context has not been extensively studied, with most research focusing on the U.S. (Albuquerque et al., 2020; Becchetti et al., 2015; Lins et al., 2017). This highlights a significant opportunity for further exploration in this domain.

Reduced downside volatility and maximum drawdown during the drawdown period experienced by sustainable portfolios compared to sin portfolios suggest that sustainable investors may have faced fewer unrealised losses during challenging periods. However, it is crucial to note that these differences may be attributed to distinct factor exposures rather than solely to sustainability criteria. Additionally, evidence of reduced volatility for sustainable portfolios outside crisis periods has been observed, suggesting that this finding may not be exclusive to such challenging times. It must be acknowledged that, due to the high tilt of sustainable portfolios towards large caps, they inherently exhibit less volatility, prompting further exploration into whether the reduction in volatility is driven by sustainability criteria or other factors.

It's crucial to acknowledge that the chosen methodology, which revolved around assessing the relationship between ESG and risk-adjusted performance during crisis periods with abnormal returns as the dependent variable, may not be the most suitable for capturing asymmetric benefits comprehensively. While the primary focus was on excess returns, it's equally essential to explore their connection with other dependent variables, such as volatility and idiosyncratic volatility, as highlighted in research by Albuquerque et al. (2020). This research revealed lower overall risk for sustainable companies during crises. Examining additional risk-adjusted measures could provide a more comprehensive understanding of the relationship during crisis periods. Unfortunately, such an exploration surpassed the scope of this thesis but presents a significant avenue for future research. Moreover, many studies investigating returns during crisis periods have adjusted for firm and/or industry characteristics, deviating from the approach of employing Fama-French risk factors, as demonstrated by works such as Albuquerque et al. (2020) and Lins et al. (2017). Alternatively, some studies, like Becchetti et al. (2015), have opted for the Sharpe ratio as a dependent variable. Neglecting the volatility factor in this study could potentially overlook crucial aspects, making it challenging to recognise the potential asymmetrical benefits associated with sustainable companies as highlighted in meta-studies (Whelan et al., 2021; Atz et al., 2023).

Finally, it is important to note that studies revealing a strong positive relationship often focused specifically on assessing the relationship during crises. These studies employed extensive screening of the portfolios under investigation, excluding certain industries, such as the financial industry (Lins et al., 2017), and selectively considering the most material factors in the ESG scores instead of taking them at face value. For instance, Lins et al. (2017) examined scores from five different ESG categories instead of the general ESG score. At the same time, Albuquerque et al. (2020) specifically considered the ES score and omitted governance aspects.

Therefore, this study did not find a direct relationship between risk-adjusted returns and ESGC scores during COVID-19, recognising the need for a more in-depth analysis focusing on overlooked volatility measures. While existing research suggests certain relationships, relying solely on general ESG information and examining only excess returns prevents a conclusive determination. In this sample, positive alphas were observed in the U.S. context and negative ones in Europe, albeit without statistical significance, indicating the possibility of these outcomes being mere chance observations in the dataset. Given the absence of statistically significant impacts throughout the entire observation period, along with some positive alphas in the recent timeframe and inconclusive findings during the COVID-19 period, conclusions are that, in terms of risk-adjusted returns, investors would have realised comparable returns, whether over extended periods or during downturn periods. This should reassure sustainable investors, indicating that there's no need for concern about potential sacrifices, as implementing sustainability primarily results in a shift in portfolio characteristics but no substantial shift in risk-adjusted performance.

### **7.3 Examining the Viability of ESG-Weighted Portfolios**

Finally, this paper explores how investors could effectively integrate ESG parameters into their portfolio allocation and whether this integration would entail additional costs. Chapter 2.1 outlined various approaches, including negative and positive screening, the best-in-class approach, and ESG integration. This study analysed the incorporation of ESG into portfolio allocation without necessarily restricting the investment universe - an issue commonly associated with traditional negative screens that may result in lower risk-adjusted returns, as indicated by numerous researchers (Geczy et al., 2005; Kempf & Osthoff, 2007; Statman & Glushkov, 2009). In alignment with Barnett and Salomon's (2006) recommendation to avoid excluding low ESG entities, this study instead incorporated ESG weights. The empirical analysis indicated that ESGC-weighted portfolios demonstrated similar risk-adjusted returns over extended periods, with slightly better performance observed in recent subperiods, consistent with the trends observed for top and bottom ESGC portfolios. This trend was evident for ESGC weighting across both equal-weighted and value-weighted portfolios, with one exception.

In the U.S. market, the value-weighted index (MKTV) outperformed the value and ESGC-weighted index (ESGV) in terms of risk-adjusted returns. This difference can be attributed, at least partially, to the impact of value weighting, where specific companies hold disproportionate representation in the portfolio, particularly in the U.S. market. Additionally, a specific ESGC score bias can be observed, where the largest companies have considerably lower ESGC scores than their counterparts due to excessive controversies. This may be explained through their sheer size, which attracts considerable media attention. Simultaneously, the top ten companies by market value, representing around 20 to 25% of overall market capitalisation, have outperformed the remaining stocks over the 14 years, with an average annualised excess return of 2.74%. Consequently, by applying ESGC and value weighting for the ESGV portfolio, their overall weight in the portfolio diminishes, leading to an overall lower performance. More

detailed data representation of this bias can be found in Appendix 6. This also underscores the rationale for analysing equal-weighted portfolios in the main analysis, aiming to overcome challenges associated with value weighting.

Except for this specific case, the other ESGC-weighted portfolios exhibited performance comparable to their benchmarks in both geographical locations, with slightly higher alphas, albeit not statistically significant. In a practical context, it is essential to account for transaction costs, as investors only realise returns after covering all associated fees. The introduction of ESGC weighting to more traditional portfolios resulted in higher annual turnover, subsequently increasing costs and reducing risk-adjusted returns. Given the overarching conclusion that markets are efficient and the incorporation of high or low ESGC companies in the portfolio doesn't result in significantly different risk-adjusted returns, the increased complexity of the portfolio leads to reduced risk-adjusted returns. The findings imply that investors should be willing to pay a very slight premium through transaction costs to rebalance the portfolio based on their preferences for social responsibility, consistent with Amon et al. (2021). However, this study provides assurance, as even with the inclusion of 200 basis points roundtrip transaction costs, alphas remained close to zero (the exception being the previously mentioned U.S. case), indicating that the impact of these additional transaction costs is relatively small in a passive portfolio management strategy.

Moreover, in the context of passive portfolio management, it is feasible to establish specific criteria aimed at partially reducing the additional turnover costs associated with ESG considerations. For instance, creating different ESG segments and allocating companies to these segments allows for over- or underweighting a traditional index based on these ESG segments without the need to reweight individual companies each time their ESG scores change by a slight percentage, as was done in this study. Additionally, considering multiple ESG ratings from different providers, opting for more stable ESG ratings and avoiding those significantly influenced by controversy scores is another viable approach.

ESG-weighted or ESG over- and underweighting indexes do exist to some extent, but their widespread adoption in ETFs is currently limited. One illustration is the MSCI ESG Universal<sup>16</sup> product panel, which adjusts the market cap-based index based on ESG scores and trends. A quick assessment of returns using their different fact sheets reveals that, on average, the ESG-weighted portfolio delivered comparable or slightly higher returns. As anticipated, this sustainability incorporation led to an increased turnover ratio, ranging from 5% to 11% across various indexes. This closely aligns with this study's observation of an 11% surge in turnover when incorporating strict ESGC weighting into value-weighted portfolios. While investors would incur slightly higher turnover costs, this study found that these costs are unlikely to significantly impact overall returns. For instance, even under the highest hypothesised roundtrip transaction costs of 200 basis points, the average annual cost increase would only be 0.22%. An illustrative instance can be drawn from Invesco, an investment management company that provides MSCI ESG

---

<sup>16</sup> *MSCI ESG Universal Indexes*. (n.d.). MSCI. <https://www.msci.com/msci-esg-universal-indexes>

Universal product panels. Specifically, their ESG Universal screened MSCI Europe ETF<sup>17</sup> incurs only a marginal 0.05% additional turnover fee compared to their conventional MSCI Europe ETF<sup>18</sup>. Therefore, in practical portfolio scenarios, the observed turnover ratios and associated costs are even lower than the initially hypothesised values.

In conclusion, ESG weighting stands out as a highly effective strategy for addressing challenges associated with basic screening methods. This approach enables passive investment approaches that yield comparable risk-adjusted returns with enhanced sustainability criteria, albeit with a slight additional premium for turnover-related fees. Furthermore, the absence of negative screening ensures that investors can have confidence in their investments, free from adverse impacts due to the exclusion of specific industries or the omission of high-performing companies.

## 7.4 Limitations & Suggestions for Further Research

In the pursuit of understanding the relationship between ESG and financial performance, the inherent limitations of the study are acknowledged. These challenges prompt an assessment of areas for improvement and further exploration. This chapter addresses the shortcomings in methodology, data sources, and analytics. Additionally, ideas for future research are suggested to enhance the understanding of the dynamics between ESG factors and investment outcomes.

Firstly, reliance on a single ESG rating may pose limitations, as incorporating a combination of multiple ratings is crucial for assessing the robustness of findings across various ESG rating providers. ESG ratings often diverge significantly (Berg et al., 2022), and a comprehensive analysis should consider multiple perspectives. Lacking access to alternative rating providers restricted this paper's capacity to incorporate a more diverse range of perspectives, posing a potential vulnerability as the results could excessively rely on Refinitiv's specific methodology. Considering the difficulty ESG rating providers face in accurately reflecting companies' sustainability, a more prudent approach might involve combining assessments from different providers, aligning with the wisdom of crowds' principle (Surowiecki, 2005). This principle emphasises that the result attains the highest accuracy when independent judgments are combined through a specific process. In the context of ESG, this translates to more precise ESG scores that effectively capture a company's sustainability. This would prove advantageous for investors and asset managers aiming to optimise the genuine sustainability profile of their portfolios. Additionally, it offers researchers the opportunity to determine the objective relationship between ESG and financial performance without being influenced by subjective opinions from various individual rating providers. A provider of such consensus ratings is OWL ESG<sup>19</sup>.

---

<sup>17</sup> Invesco MSCI Europe ESG Universal Screened UCITS ETF Acc, ISIN : IE00BJQRDL90

<sup>18</sup> Invesco MSCI Europe UCITS ETF Acc, ISIN : IE00B60SWY32

<sup>19</sup> OWL ESG. (n.d.). *OWL ESG Consensus Scores*. Retrieved March 10, 2024, from <https://owlesg.com/consensus-scores/>

Secondly, a specific bias in the utilised ESGC score was observed. The bias consists in the phenomenon where the largest companies had lower ESGC scores than average companies (see Appendix 6), possibly attributed to heightened media scrutiny. This bias compounds with the well-established bias that smaller firms typically receive lower simple ESG ratings, attributed to their limited resources and lesser pressure to disclose comprehensive ESG data (Sipiczki, 2022; Dremptec et al., 2020). In this context, the utilised ESGC had lower ratings for both smaller and the largest companies. This is primarily attributed to the largest companies facing a significantly higher number of controversies compared to regular large companies, owing to increased media scrutiny, even with the existing procedure Refinitiv (2022) has implemented to address size bias in the ESGC score. To address this challenge, future studies could consider normalising ESG and ESGC scores relative to companies' market capitalisation to mitigate market cap bias.

A third potential limitation of this study could arise from portfolios displaying dynamic characteristics, as evident in Table 3, where a certain trend in the average market capitalisation of low ESGC portfolios compared to the overall portfolio over time was observed. Thus, it is conceivable that other portfolio characteristics, such as value, profitability, and investment, also exhibited certain dynamic features over time. The annual rebalancing of portfolios based on ESGC scores and changes in constituent lists resulted in stock turnover, potentially leading to changes in portfolio characteristics over time. Given that the utilised multi-factor regression models account for portfolio characteristics by controlling for these risk factors, notable fluctuations may yield unreliable outcomes, potentially compromising the accuracy of identifying return drivers. Thus, it may be interesting to explore other regression methodologies to accommodate dynamic parameters, such as rolling time-series analysis (Zivot & Wang, 2003).

Lastly, this study did not entirely account for survivorship bias, which arises when the analysis includes data only from companies that have survived or remained in the dataset until the end of the observation period (Brown et al., 1992). In the approach adopted, companies were included for which both return and ESGC data were available for an entire year to construct the yearly portfolios. This exclusion of companies that may have ceased to exist or were delisted could impact the representativeness of the findings, potentially leading to an overestimation of performance as struggling companies are omitted. If struggling companies are predominantly low or high ESGC performers, this could result in inaccurate conclusions. However, the applied methodology, involving the construction of new portfolios every year rather than relying solely on one portfolio to cover the 14-year observation period, should help to mitigate survivorship bias to some extent.

In addition to outlining the limitations, two specific avenues for further research are proposed, directly stemming from the findings presented in this paper.

A promising direction for future research lies in exploring the post-COVID-19 era, considering recent geopolitical events such as the conflict in Ukraine and the inflation surge across developed nations. This is motivated by the discovery of positive alphas in the subperiod from 2017 to 2023, where the absence of significant impacts from the

COVID-19 pandemic, both during the drawdown and subsequent recovery period, suggests alternative drivers for these positive alphas. Moreover, the empirical analysis revealed a sharpening return divergence from 2021 onwards. Given that at least two other studies have reported positive alphas in this recent subperiod (De Spiegeleer et al., 2023; Teti et al., 2023), further investigation into this phenomenon, its robustness across different sustainability ratings, and its driving factors would be intriguing.

Another avenue closely related to the broader outcomes of this study, which indicated the absence of statistically significant differences in returns over extended periods (2010 – 2023) and the overall efficiency of markets, involves exploring the impact of various ESG integration strategies on annual management and turnover costs within Exchange Traded Funds (ETFs). The goal is to determine the best equilibrium between improving sustainability and reducing related expenses, all while ensuring a diverse investment universe. This investigation holds the potential to inform the development of innovative ESG integration strategies in passively managed funds, presenting investors with advanced alternatives beyond the commonly employed screening criteria for constructing sustainable ETFs, thereby avoiding a substantial reduction in the investment universe. Such a study would immensely benefit asset management firms seeking creative ways to optimise their ESG integration strategy. The insights gained could lead to creating more competitive and appealing investment products. This would also be advantageous for investors seeking passive investment options that integrate ESG criteria without imposing constraints on the available investment opportunities. This approach aims to offer comparable risk-adjusted returns while minimising associated fees, presenting an attractive proposition for investors seeking sustainability without compromising financial performance.

## 8 Conclusion

This study investigated the relationship between ESG considerations and risk-adjusted financial performance from investors' perspective. Employing a quantitative analysis, diverse portfolios were constructed using constituents from the S&P 500 and STOXX Europe 600 indices based on Refinitiv ESG combined scores. Findings indicated that there were no substantial or consistent variations in risk-adjusted returns across high ESGC, low ESGC, or ESGC-weighted portfolios when examined through the lens of the Fama-French 5-factor or Carhart 4-factor model, suggesting that over extended periods, the market is efficient. Despite discovering statistically significant positive alphas in recent years, their robustness was lacking in different multi-factor regressions or value-weighted portfolios. While deviations from market efficiency might be plausible in recent years due to evolving landscapes and behaviours, this finding could also be attributed to certain limitations in this study. Additionally, amid the COVID-19 pandemic, no statistically significant alpha variations were discerned. However, during the drawdown period, sustainable companies in the U.S. exhibited higher alphas, contrasting with lower alphas in Europe. In essence, investors would not consistently encounter significant advantages or penalties over long periods by choosing more sustainable portfolios. The primary distinction lies in the exposure of their portfolios to other factors, such as large-cap value stocks, and reduced susceptibility to market risk. Additionally, no significant differences in risk-adjusted returns emerged when accounting for potential increased turnover costs associated with ESGC-weighted portfolios. This implies that, despite a marginal uptick in turnover-related fees, the impact is not substantial enough to significantly disadvantage sustainable investors, resulting in overall comparable risk-adjusted returns. While this conclusion holds for all equal-weighted portfolios, variations in findings were evident for value-weighted portfolios in the U.S., indicating a negative relationship. This could be attributed to lower diversification and a specific ESGC score bias.

Hence, the response to the research question indicates that incorporating ESG data in the construction of stock portfolios does not significantly impact risk-adjusted returns over extended periods. Consequently, investors cannot exploit this information to enhance their risk-adjusted returns. However, they can use this information to increase the average ESG score of their portfolio, such as through ESG-weighted portfolios, thereby enhancing overall sustainability without sacrificing financial performance - an approach aligned with the principle of doing good while doing well.

Additionally, several directions for further research are proposed, particularly emphasising the need for more precise examinations of the relationship during crisis periods in Europe, where existing studies are limited, and exploring the post-COVID-19 period to further scrutinise the potential positive alpha. Moreover, exploring the relationship based on objective consensus ESG scores presents an intriguing research direction, addressing challenges that many studies encounter due to their absence of data from multiple ESG rating providers. Finally, conducting more thorough examinations into the integration of ESG weighting into Exchange Traded Funds (ETFs) and the anticipated

returns for investors is strongly recommended. This is of utmost importance, given the limited research in this domain and considering that ETFs are a favoured choice among investors aiming for diversified portfolios with minimal costs. While numerous ETFs currently incorporate sustainability criteria, the majority employ negative screening. Further exploration into the impact of weighting ETFs based on companies' sustainability would provide investors with enhanced diversification opportunities at reduced costs without overlooking potential profitable investments. Consequently, a similar study could lay the foundation for a potential new approach to sustainable passive investment, benefiting both asset managers pioneering such portfolios and investors seeking diverse methods to incorporate sustainability criteria.

As a concluding remark, investors can confidently incorporate sustainability into their investment decisions without adversely impacting returns. However, it is essential to demonstrate the availability of easy and cost-effective methods for seamless integration, emphasising the significance of passive-managed ETFs. While this study affirms the compatibility of sustainability with returns, it remains crucial to showcase accessible and cost-effective approaches. The rapidly growing segment of ESG ETFs, surpassing their actively managed counterparts in the U.S., opens doors for extensive research and innovation (KPMG, 2022). The increasing risk of oversaturation intensifies competition for product differentiation, spotlighting inventive strategies such as ESG weighting.

## References

- Abate, G., Basile, I., & Ferrari, P. (2021). The level of sustainability and mutual fund performance in Europe: An empirical analysis using ESG ratings. *Corporate Social Responsibility and Environmental Management*, 28(5), 1446-1455.
- Albuquerque, R., Koskinen, Y., & Zhang, C. (2019). Corporate social responsibility and firm risk: Theory and empirical evidence. *Management Science*, 65(10), 4451-4469.
- Albuquerque, R., Koskinen, Y., Yang, S., & Zhang, C. (2020). Resiliency of environmental and social stocks: An analysis of the exogenous COVID-19 market crash. *The Review of Corporate Finance Studies*, 9(3), 593-621.
- Alley, I., Adebayo, A., & Oligbi, B. (2016). Corporate Governance and Financial Performance Nexus: Any Bidirectional Causality? *International Journal of Management and Economics*, 50(1), 82-99.
- Allouche, J., & Laroche, P. (2005). A meta-analytical investigation of the relationship between corporate social and financial performance. *Revue de gestion des ressources humaines*, 57, 18.
- Aluchna, M., Roszkowska-Menkes, M., & Kamiński, B. (2022). From talk to action: the effects of the non-financial reporting directive on ESG performance. *Meditari Accountancy Research*, 31(7), 1-25.
- 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.
- Amon, J., Rammerstorfer, M., & Weinmayer, K. (2021). Passive ESG portfolio management—the benchmark strategy for socially responsible investors. *Sustainability*, 13(16), 9388.
- Antonicic, M. (2019). Why sustainability? Because risk evolves and risk management should too. *Journal of Risk Management in Financial Institutions*, 12(3), 206-216.
- Atz, U., Van Holt, T., Liu, Z. Z., & Bruno, C. C. (2023). Does sustainability generate better financial performance? review, meta-analysis, and propositions. *Journal of Sustainable Finance & Investment*, 13(1), 802-825.
- Bannier, C. E., Bofinger, Y., & Rock, B. (2019). Doing safe by doing good: ESG investing and corporate social responsibility in the US and Europe. *CFS Working Paper Series*, 621.
- Barnett, M. L., & Salomon, R. M. (2006). Beyond dichotomy: The curvilinear relationship between social responsibility and financial performance. *Strategic Management Journal*, 27(11), 1101-1122.
- Barnett, M. L., & Salomon, R. M. (2012). Does it pay to be really good? Addressing the shape of the relationship between social and financial performance. *Strategic Management Journal*, 33(11), 1304-1320.
- Bauer, R., Koedijk, K., & Otten, R. (2005). International evidence on ethical mutual fund performance and investment style. *Journal of Banking & Finance*, 29(7), 1751-1767.
- Becchetti, L., Ciciretti, R., Dalò, A., & Herzel, S. (2015). Socially responsible and conventional investment funds: performance comparison and the global financial crisis. *Applied Economics*, 47(25), 2541-2562.
- Becker, M. G., Martin, F., & Walter, A. (2022). The power of ESG transparency: The effect of the new SFDR sustainability labels on mutual funds and individual investors. *Finance Research Letters*, 47, 102708.
- Bello, Z. Y. (2005). Socially responsible investing and portfolio diversification. *Journal of Financial Research*, 28(1), 41-57.
- Bello, Z. Y. (2008). A statistical comparison of the CAPM to the Fama-French Three Factor Model and the Carhart's Model. *Global Journal of Finance and Banking Issues*, 2(2).
- Berg, F., Koelbel, J. F., & Rigobon, R. (2022). Aggregate confusion: The divergence of ESG ratings. *Review of Finance*, 26(6), 1315-1344.
- Bhagat, S., & Bolton, B. (2008). Corporate governance and firm performance. *Journal of Corporate Finance*, 14(3), 257-273.
- Bhagat, S., & Bolton, B. (2019). Corporate governance and firm performance: The sequel. *Journal of Corporate Finance*, 58, 142-168.
- Blanco, B. (2012). The use of CAPM and Fama and French Three Factor Model: portfolios selection. *Public and Municipal Finance*, 1(2), 61-70.

- Bodie, Z., Kane, A., & Marcus, A. J. (2018). *Investments: Tenth Edition*. New York: McGraw-Hill Education.
- Bollen, N. P. (2007). Mutual fund attributes and investor behavior. *Journal of financial and quantitative analysis*, 42(3), 683-708.
- Broadstock, D. C., Chan, K., Cheng, L. T., & Wang, X. (2021). The role of ESG performance during times of financial crisis: Evidence from COVID-19 in China. *Finance Research Letters*, 38, 101716.
- Brown, S. J., Goetzmann, W., Ibbotson, R. G., & Ross, S. A. (1992). Survivorship bias in performance studies. *The Review of Financial Studies*, 6(4), 553-580.
- Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, 52(1), 57-82.
- 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.
- Clark, G. L., & Viehs, M. (2014). The implications of corporate social responsibility for investors: An overview and evaluation of the existing CSR literature. Available at SSRN 2481877.
- Conover, C., Jensen, G., Johnson, R., & Mercer, J. (2005). Is Fed Policy Still Relevant for Investors? *Financial Analysts Journal*, 61, 70-79.
- Das, N., Chatterjee, S., Sunder, A., & Ruf, B. (2018). ESG ratings and the performance of socially responsible mutual funds: A panel study. *Journal of Finance Issues*, 17(1), 49-57.
- de Quevedo-Puente, E., De La Fuente-Sabaté, J. M., & Delgado-García, J. B. (2007). Corporate social performance and corporate reputation: Two interwoven perspectives. *Corporate Reputation Review*, 10, 60-72.
- De Spiegeleer, J., Höcht, S., Jakubowski, D., Reyners, S., & Schoutens, W. (2023). ESG: A new dimension in portfolio allocation. *Journal of Sustainable Finance & Investment*, 13(2), 827-867.
- Del Giudice, A., & Rigamonti, S. (2020). Does audit improve the quality of ESG scores? Evidence from corporate misconduct. *Sustainability*, 12(14), 12(14), 5670.
- Derwall, J., Guenster, N., Bauer, R., & Koedijk, K. (2005). The eco-efficiency premium puzzle. *Financial Analysts Journal*, 61(2), 51-63.
- Dimson, E., Marsh, P., & Staunton, M. (2020). Divergent ESG ratings. *The Journal of Portfolio Management*, 47(1), 75-87.
- Directive 2014/95. (n.d.). *Directive 2014/95/EU of the European Parliament and of the Council of 22 October 2014 amending Directive 2013/34/EU as regards disclosure of non-financial and diversity information by certain large undertakings and groups*. <http://data.europa.eu/eli/dir/2014/95/oj>.
- Directive 2022/2464. (n.d.). *Directive (EU) 2022/2464 of the European Parliament and of the Council of 14 December 2022 amending Regulation (EU) No 537/2014, Directive 2004/109/EC, Directive 2006/43/EC and Directive 2013/34/EU, as regards corporate sustainability reporting*. <http://data.europa.eu/eli/dir/2022/2464/oj>.
- Dobrick, J., Klein, C., & Zwergel, B. (2023). Size bias in refinitiv ESG data. *Finance Research Letters*, 55, 104014.
- Doyle, T. (2018). *Ratings That Don't Rate: The Subjective World of ESG Ratings Agencies*. American Council for Capital Formation.
- Drempetic, S., Klein, C., & Zwergel, B. (2020). The Influence of Firm Size on the ESG Score: Corporate Sustainability Ratings Under Review. *Journal of Business Ethics*, 167, 333-360.
- Dumrose, M., Rink, S., & Eckert, J. (2022). Disaggregating confusion? The EU Taxonomy and its relation to ESG rating. *Finance Research Letters*, 48, 102928.
- Endrikat, J., Guenther, E., & Hoppe, H. (2014). Making sense of conflicting empirical findings: A meta-analytic review of the relationship between corporate environmental and financial performance. *European Management Journal*, 32(5), 735-751.
- Engle, R. F., Giglio, S., Kelly, B., Lee, H., & Stroebe, J. (2020). Hedging climate change news. *The Review of Financial Studies*, 33(3), 1184-1216.
- ESMA. (2022a). *Call for evidence On Market Characteristics for ESG Rating Providers in the EU*.
- ESMA. (2022b). *The drivers of the costs and performance of ESG funds*.
- Esty, D. C., & Porter, M. E. (1998). Industrial ecology and competitiveness: Strategic implications for the firm. *Journal of Industrial Ecology*, 2(1), 35-43.
- Eurosif. (2018). *European SRI Study 2018*.

- Fabozzi, F. J., Gupta, F., & Markowitz, H. M. (2002). The legacy of modern portfolio theory. *The Journal of Investing*, 11(3), 7-22.
- Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25(2), 383-417.
- Fama, E. F. (1998). Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics*, 49(3), 283-306.
- Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1), 3-56.
- Fama, E. F., & French, K. R. (2007). Disagreement, tastes, and asset prices. *Journal of Financial Economics*, 83(3), 667-689.
- Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1), 1-22.
- Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, 81(3), 607-636.
- Folqué, M., Escrig-Olmedo, E., & Corzo Santamaría, T. (2021). Sustainable development and financial system: Integrating ESG risks through sustainable investment strategies in a climate change context. *Sustainable Development*.
- Frazier, L. (2021, February 11). *The Coronavirus Crash Of 2020, And The Investing Lesson It Taught Us*. Retrieved February 2024, from Forbes: <https://www.forbes.com/sites/lizfrazierpeck/2021/02/11/the-coronavirus-crash-of-2020-and-the-investing-lesson-it-taught-us/?sh=58dba26646cf>
- Frecautan, I., & Nita, A. (2022). WHO IS GOING TO WIN: THE EU ESG REGULATION OR THE REST OF THE WORLD?—A CRITICAL REVIEW. *Annals of Faculty of Economics*, 2(2), 109-120.
- Freeman, R. E. (2010). Strategic management: A stakeholder approach. *Cambridge University Press*.
- 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.
- Friedman, M. (1970, September 13). The Social Responsibility of Business Is to Increase Its Profits. *New York Times Magazine*, 122-126.
- Geczy, C., Stambaugh, R. F., & Levin, D. (2005). Investing in socially responsible mutual funds. Available at SSRN 416380.
- Giglio, S., Maggiori, M., Stroebel, J., Tan, Z., Utkus, S., & Xu, X. (2023). Four facts about ESG beliefs and investor portfolios. *National Bureau of Economic Research*, 31114.
- GSIA. (2020). *Global Sustainable Investment Review 2020*.
- GSIA. (2022). *Global Sustainable Investment Review 2022*.
- Halbritter, G., & Dorfleitner, G. (2015). The wages of social responsibility—where are they? A critical review of ESG investing. *Review of Financial Economics*, 26, 25-35.
- Hang, M., Geyer-Klingenberg, J., & Rathgeber, A. W. (2019). It is merely a matter of time: A meta-analysis of the causality between environmental performance and financial performance. *Business Strategy and the Environment*, 28(2), 257-273.
- Hart, S. L. (1995). A natural-resource-based view of the firm. *Academy of Management Review*, 20(4), 986-1014.
- Herzel, S., Nicolosi, M., & Stărică, C. (2012). The cost of sustainability in optimal portfolio decisions. *The European Journal of Finance*, 18(3-4), 333-349.
- Humphrey, J. E., Lee, D. D., & Shen, Y. (2012). Does it cost to be sustainable? *Journal of Corporate Finance*, 18(3), 626-639.
- Hvidkjær, S. (2017). ESG investing: a literature review. *Report prepared for Dansif*.
- IFAC. (2022). *The State of Play in Reporting and Assurance of Sustainability Information: Update 2019-2020 Data & Analysis*.
- Jamali, D., Safieddine, A. M., & Rabbath, M. (2008). Corporate Governance and Corporate Social Responsibility Synergies and Interrelationships. *Corporate governance: an international review*, 16(5), 443-459.
- Jensen, M. C. (1968). The performance of mutual funds in the period 1945-1964. *The Journal of Finance*, 23(2), 389-416.
- Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3(4), 305-360.

- Jin, I. (2022). ESG-screening and factor-risk-adjusted performance: The concentration level of screening does matter. *Journal of Sustainable Finance & Investment*, 12(4), 1125-1145.
- Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis Decision Under Risk. *Econometrica*, 46(2), 263-292.
- Kempf, A., & Osthoff, P. (2007). The effect of socially responsible investing on portfolio performance. *European Financial Management*, 13(5), 908-922.
- KPMG. (2022). *Evolution of ESG investing: Near-term implications and strategies for asset managers*.
- Krosinsky, C. (2014). The long and necessary death of socially responsible investing. *Journal of Sustainable Finance & Investment*, 4(3), 297-298.
- Krueger, P., Sautner, Z., Tang, D. Y., & Zhong, R. (2021). The effects of mandatory ESG disclosure around the world. *European Corporate Governance Institute-Finance Working Paper*(754), 21-44.
- Larcker, D. F., Pomorski, L., Tayan, B., & Watts, E. M. (2022). ESG ratings: A Compass without direction. *Rock Center for Corporate Governance at Stanford University Working Paper Forthcoming*.
- Lee, D. D., Humphrey, J. E., Benson, K. L., & Ahn, J. Y. (2010). Socially responsible investment fund performance: the impact of screening intensity. *Accounting & Finance*, 50(2), 351-370.
- Levitt, T. (1958). The dangers of social-responsibility. *Harvard Business Review*, 36(5), 41-50.
- Lins, K. V., Servaes, H., & Tamayo, A. (2017). Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis. *The Journal of Finance*, 72(4), 1785-1824.
- Lintner, J. (1965). Security prices, risk, and maximal gains from diversification. *The Journal of Finance*, 20(4), 587-615.
- Liu, W., & Strong, N. (2008). Biases in decomposing holding-period portfolio returns. *The Review of Financial Studies*, 21(5), 2243-2274.
- Magdon-Ismael, M., & Atiya, A. F. (2004). Maximum drawdown. *Risk Magazine*, 17(10), 99-102.
- Malkiel, B. G. (2003). The efficient market hypothesis and its critics. *Journal of Economic Perspectives*, 17(1), 59-82.
- Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, 7(1), 77-91.
- Milonas, N., Rompotis, G., & Moutzouris, C. (2022). The performance of ESG funds vis-à-vis non-ESG funds. *The Journal of Impact and ESG Investing*, 2(4), 96-115.
- Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica: Journal of the Econometric Society*, 768-783.
- Orlitzky, M., Schmidt, F. L., & Rynes, S. L. (2003). Corporate social and financial performance: A meta-analysis. *Organization studies*, 24(3), 403-441.
- Paniagua, J., Rivelles, R., & Sapena, J. (2018). Corporate governance and financial performance: The role of ownership and board structure. *Journal of Business Research*, 89, 229-234.
- Pedersen, L. H., Fitzgibbons, S., & Pomorski, L. (2021). Responsible investing: The ESG-efficient frontier. *Journal of Financial Economics*, 142(2), 572-597.
- Petridis, K., Kiosses, N., Tampakoudis, I., & Ben Abdelaziz, F. (2023). Measuring the efficiency of mutual funds: Does ESG controversies score affect the mutual fund performance during the COVID-19 pandemic? *Operational Research*, 23(3), 54.
- Platform on Sustainable Finance. (2022). *Final Report on Social Taxonomy*.
- Preston, L. E., & O'Bannon, D. P. (1997). The corporate social-financial performance relationship: A typology and analysis. *Business & Society*, 36(4), 419-429.
- Qi, Y., & Li, X. (2020). On imposing ESG constraints of portfolio selection for sustainable investment and comparing the efficient frontiers in the weight space. *Sage Open*, 10(4).
- Redondo Alamillos, R., & de Mariz, F. (2022). How can European regulation on ESG impact business globally? *Journal of Risk and Financial Management*, 15(7), 291.
- Refinitiv. (2022). *Environmental, social and governance scores from Refinitiv*.
- Regulation 2019/2088. (n.d.). *REGULATION (EU) 2019/2088 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 November 2019 on sustainability-related disclosures in the financial services sector*. <http://data.europa.eu/eli/reg/2019/2088/oj>.
- Rehnby, N. (2016). Does the Fama-French three-factor model and Carhart four-factor model explain portfolio returns better than CAPM. A study performed on the Swedish stock market.

- Renneboog, L., Ter Horst, J., & Zhang, C. (2008). The price of ethics and stakeholder governance: The performance of socially responsible mutual funds. *Journal of Corporate Finance*, 14(3), 302-322.
- Renneboog, L., Ter Horst, J., & Zhang, C. (2011). Is ethical money financially smart? Nonfinancial attributes and money flows of socially responsible investment funds. *Journal of Financial Intermediation*, 20(4), 562-588.
- Rollinger, T. N., & Hoffman, S. T. (2013). *Sortino: a 'sharper' ratio*. Chicago, Illinois: Red Rock Capital.
- Rossi, M., Nerino, M., & Capasso, A. (2015). Corporate governance and financial performance of Italian listed firms. The results of an empirical research. *Corporate Ownership & Control*, 12(2), 628-643.
- Sahin, Ö., Bax, K., Czado, C., & Paterlini, S. (2022). Environmental, Social, Governance scores and the Missing pillar—Why does missing information matter?. *Corporate Social Responsibility and Environmental Management*, 29(5), 1782-1798.
- Sahut, J. M., Peris-Ortiz, M., & Teulon, F. (2019). Corporate social responsibility and governance. *Journal of Management and Governance*, 23, 901-912.
- Schueth, S. (2003). Socially responsible investing in the United States. *Journal of Business Ethics*, 43, 189-194.
- SEC 33-11042/34-94478. (n.d.). *Proposed Rule: The Enhancement and Standardization of Climate-Related Disclosures for Investors*. <https://www.federalregister.gov/d/2022-06342>.
- Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19(3), 425-442.
- Sinha, R., Datta, M., & Ziolo, M. (2019). Inclusion of ESG Factors in Investments and Value Addition: A Meta-Analysis of the Relationship. In *Effective Investments on Capital Markets: 10th Capital Market Effective Investments Conference (CMEI 2018)*, 93-109.
- Sipiczki, A. (2022). *A critical look at the ESG market*. CEPS.
- Statman, M., & Glushkov, D. (2009). The wages of social responsibility. *Financial Analysts Journal*, 65(4), 33-46.
- Steen, M., Moussawi, J. T., & Gjølberg, O. (2020). Is there a relationship between Morningstar's ESG ratings and mutual fund performance? *Journal of Sustainable Finance & Investment*, 10(4), 349-370.
- Stock, J. H., & Watson, M. W. (2020). *Introduction to Econometrics (4th Edition)*. Pearson.
- Surowiecki, J. (2005). *The wisdom of crowds*. Anchor.
- Tamimi, N., & Sebastianelli, R. (2017). Transparency among S&P 500 companies: An analysis of ESG disclosure scores. *Management Decision*, 55(8), 1660-1680.
- Teti, E., Dallochio, M., & L'Erario, G. (2023). The impact of ESG tilting on the performance of stock portfolios in times of crisis. *Finance Research Letters*, 52, 103522.
- Trumpp, C., & Guenther, T. (2017). Too little or too much? Exploring U-shaped relationships between corporate environmental performance and corporate financial performance. *Business Strategy and the Environment*, 26(1), 49-68.
- Waddock, S. A., & Graves, S. B. (1997). The corporate social performance–financial performance link. *Strategic Management Journal*, 18(4), 303-319.
- Whelan, T., Atz, U., & Clark, C. (2021). ESG and Financial Performance: Uncovering the Relationship by Aggregating Evidence from 1,000 Plus Studies.
- Wimmer, M. (2013). ESG-persistence in socially responsible mutual funds. *J. Mgmt. & Sustainability*, 3(9).
- Zivot, E., & Wang, J. (2003). Rolling analysis of time series. In E. Zivot, & J. Wang, *Modeling financial time series with S-Plus®* (pp. 299-346).
- Zumante, I., & Lăce, N. (2021). ESG Rating—Necessity for the Investor or the Company? *Sustainability*.

# Appendices

## Appendix 1: Statistical Tests

Appendix 1 presents results from various statistical tests, such as the Augmented Dickey-Fuller test (Appendix 1.1), Durbin Watson test (Appendix 1.2), and White test (Appendix 1.3). The below appendices include outcomes from the main portfolios and regressions, excluding those from robustness tests or those involving transaction costs. Nevertheless, after conducting tests on all portfolios and regressions, the results remain consistent: all portfolios demonstrate stationarity, with few having some levels of autocorrelation and most having high levels of heteroscedasticity.

### Appendix 1.1: Testing for Stationarity

Tables 18 and 19 showcase the results of the Augmented Dickey-Fuller unit root test on the S&P 500 and STOXX Europe portfolios, respectively. The null hypothesis states that the time series are non-stationary, while the alternative hypothesis suggests they are stationary. As indicated in the tables below, the p-values for all portfolios are less than 1%, leading to the rejection of the null hypothesis at a 1% confidence level and the conclusion that all portfolios in both indices are stationary.

**Table 18 : Augmented Dickey-Fuller test results for portfolios based on S&P 500 constituents**

	T20E	B20E	L-S	ESGE	ESGV	MKTE	MKTV	Obs.
2010 - 2023	-14.856*** (0.0000)	-14.619*** (0.0000)	-12.237*** (0.0000)	-14.815*** (0.0000)	-15.382*** (0.0000)	-14.789*** (0.0000)	-15.372*** (0.0000)	168
2010 - 2016	-10.313*** (0.0000)	-10.285*** (0.0000)	-8.696*** (0.0000)	-10.229*** (0.0000)	-10.399*** (0.0000)	-10.225*** (0.0000)	-10.366*** (0.0000)	84
2017 - 2023	-10.532*** (0.0000)	-10.315*** (0.0000)	-8.743*** (0.0000)	-10.547*** (0.0000)	-11.090*** (0.0000)	-10.527*** (0.0000)	-11.095*** (0.0000)	84
COVID-19	-14.897*** (0.0000)	-14.389*** (0.0000)	-11.199*** (0.0000)					135
COVID-19 drawdown	-8.317*** (0.0000)	-8.094*** (0.0000)	-5.171*** (0.0000)					34
COVID-19 recovery	-9.960*** (0.0000)	-10.051*** (0.0000)	-10.686*** (0.0000)					101

Note: This table illustrates the Dickey-Fuller test statistics for portfolios constructed using constituents from the S&P 500 as the underlying basis. Levels of statistical significance are denoted by \*\*\* (1%), \*\* (5%), and \* (10%).

**Table 19 : Augmented Dickey-Fuller test results for portfolios based on STOXX Europe 600 constituents**

	T20E	B20E	L-S	ESGE	ESGV	MKTE	MKTV	Obs.
2010 - 2023	-13.415*** (0.0000)	-13.481*** (0.0000)	-15.423*** (0.0000)	-13.493*** (0.0000)	-13.891*** (0.0000)	-13.493*** (0.0000)	-13.887*** (0.0000)	168
2010 - 2016	-9.460*** (0.0000)	-9.748*** (0.0000)	-11.545*** (0.0000)	-9.587*** (0.0000)	-9.837*** (0.0000)	-9.615*** (0.0000)	-9.874*** (0.0000)	84
2017 - 2023	-9.404 *** (0.0000)	-9.313*** (0.0000)	-10.177*** (0.0000)	-9.405*** (0.0000)	-9.702*** (0.0000)	-9.387*** (0.0000)	-9.668*** (0.0000)	84
COVID-19	-10.757*** (0.0000)	-10.266*** (0.0000)	-13.713*** (0.0000)					138
COVID-19 drawdown	-5.268*** (0.0000)	-4.962*** (0.0000)	-7.543*** (0.0000)					34
COVID-19 recovery	-9.953*** (0.0000)	-9.693*** (0.0000)	-10.540*** (0.0000)					104

Note: This table illustrates the Dickey-Fuller test statistics for portfolios constructed using constituents from the STOXX Europe 600 as the underlying basis. Levels of statistical significance are denoted by \*\*\* (1%), \*\* (5%), and \* (10%).

## Appendix 1.2: Testing for Autocorrelation

Tables 20 and 21 display the outcomes of the Durbin-Watson test for autocorrelation conducted on the Fama-French 5-factor regression applied to portfolios derived from the S&P 500 and STOXX Europe 600 indices, respectively. The d statistic is employed to identify the presence (and the sign) or absence of autocorrelation. A d statistic below 2 indicates positive autocorrelation, while a d statistic above 2 suggests negative autocorrelation. Determining statistical significance hinges on comparing the d statistic to the lower critical value dl and the upper critical value du (in this case, at a 5% confidence level). Statistically significant autocorrelation is confirmed when the observed d value is lower than dl for positive autocorrelation or when (4-d) is lower than dl for negative autocorrelation. No autocorrelation is present if the observed d value or (4-d) exceeds du. If they fall between dl and du, the test outcome is inconclusive. Results indicate that most portfolios are not serially correlated, but exceptions do arise, with some portfolios' autocorrelation being inconclusive, some statistically significantly positive and some statistically significantly negative. During the COVID-19 drawdown period, most portfolios display a positive correlation, which aligns with the scenario where portfolios experienced consecutive days of negative returns. In the broader observation period, there is a tendency for the S&P 500 index to exhibit a positive correlation, while the STOXX Europe 600 index displays a negative correlation. However, most data is not statistically significant.

**Table 20 :** *Durbin-Watson test results for portfolios based on S&P 500 constituents*

	T20E	B20E	L-S	ESGE	ESGV	MKTE	MKTV	dl	du
2010 - 2023	1.8700	1.8637	1.9215	1.9367	1.8687	1.9544	1.7842	1.623	1.725
2010 - 2016	1.7860	2.0947	1.9729	1.9022	1.9600	2.0049	2.3025	1.386	1.630
2017 - 2023	1.9724	1.8624	1.9255	2.0533	1.8880	2.0476	1.6680	1.386	1.630
COVID-19	1.3394	1.5629	1.4805					1.482	1.689
COVID-19 drawdown	1.2150	1.1026	1.1224					0.869	1.677
COVID-19 recovery	1.4882	2.0521	1.7758					1.421	1.670

Note: This table illustrates the Durbin-Watson test statistics for portfolios constructed using constituents from the S&P 500 as the underlying basis, as well as the critical values dl and du at a 5% confidence interval.

**Table 21 :** *Durbin-Watson test results for portfolios based on STOXX Europe 600 constituents*

	T20E	B20E	L-S	ESGE	ESGV	MKTE	MKTV	dl	du
2010 - 2023	2.1965	2.2195	2.2489	2.3112	2.5144	2.3063	2.4783	1.623	1.725
2010 - 2016	2.0963	2.2494	2.2798	2.2912	2.3495	2.3114	2.4166	1.386	1.630
2017 - 2023	2.0839	2.3332	2.3735	2.1932	2.2965	2.2269	2.2319	1.386	1.630
COVID-19	1.8567	1.8425	1.7835					1.482	1.689
COVID-19 drawdown	1.1639	1.1972	1.1843					0.869	1.677
COVID-19 recovery	2.2632	2.3368	1.6463					1.421	1.670

Note: This table illustrates the Durbin-Watson test statistics for portfolios constructed using constituents from the STOXX Europe 600 as the underlying basis, as well as the critical values dl and du at a 5% confidence interval.

### Appendix 1.3: Testing for Heteroscedasticity

Tables 22 and 23 present the results of the White test for heteroscedasticity conducted on the Fama-French 5-factor regression for portfolios derived from the S&P 500 and STOXX Europe 600 indices, respectively. The null hypothesis posits homoscedasticity, while the alternative hypothesis suggests unrestricted heteroscedasticity. Results indicate that the null hypothesis can be rejected for most regressions, concluding that most of the regressions exhibit heteroscedasticity.

**Table 22 :** White test results for portfolios based on S&P 500 constituents

	T20E	B20E	L-S	ESGE	ESGV	MKTE	MKTV	df
2010 - 2023	50.19*** (0.0002)	67.66*** (0.0000)	21.26 (0.3822)	68.06*** (0.0000)	54.33*** (0.0001)	71.45*** (0.0000)	68.65*** (0.0000)	20
2010 - 2016	26.50 (0.1499)	43.07*** (0.0020)	17.32 (0.6318)	29.73* (0.0744)	37.84*** (0.0093)	29.88*** (0.0718)	48.81*** (0.0003)	20
2017 - 2023	24.58 (0.2179)	36.69** (0.0127)	21.40 (0.3742)	39.89*** (0.0052)	38.35*** (0.0080)	41.30*** (0.0034)	42.99*** (0.0020)	20
COVID-19	55.41*** (0.0000)	59.50*** (0.0000)	35.67** (0.0168)					20
COVID-19 drawdown	32.48** (0.0384)	27.91 (0.1116)	27.13 (0.1317)					20
COVID-19 recovery	27.65 (0.1179)	26.89 (0.1384)	13.83 (0.8391)					20

Note: This table illustrates the White test statistics for portfolios constructed using constituents from the S&P 500 as the underlying basis. Levels of statistical significance are denoted by \*\*\* (1%), \*\* (5%), and \* (10%).

**Table 23 :** White test results for portfolios based on STOXX Europe 600 constituents

	T20E	B20E	L-S	ESGE	ESGV	MKTE	MKTV	df
2010 - 2023	53.79*** (0.0001)	21.17 (0.3872)	18.13 (0.5786)	56.40*** (0.0000)	67.22*** (0.0000)	53.36*** (0.0001)	73.66*** (0.0000)	20
2010 - 2016	41.17*** (0.0035)	30.74** (0.0587)	25.26 (0.1917)	41.95*** (0.0028)	19.06 (0.5178)	40.87*** (0.0039)	17.51 (0.6199)	20
2017 - 2023	46.99*** (0.0006)	19.89 (0.4647)	31.75** (0.0460)	27.85 (0.1130)	52.93*** (0.0001)	22.78 (0.2995)	55.19*** (0.0000)	20
COVID-19	42.85*** (0.0021)	58.15*** (0.0000)	96.87*** (0.0000)					20
COVID-19 drawdown	16.05 (0.7137)	23.00 (0.2886)	30.96* (0.0557)					20
COVID-19 recovery	42.02*** (0.0028)	22.91 (0.2935)	31.67* (0.0469)					20

Note: This table illustrates the White test statistics for portfolios constructed using constituents from the STOXX Europe 600 as the underlying basis. Levels of statistical significance are denoted by \*\*\* (1%), \*\* (5%), and \* (10%).

## Appendix 2: Robustness Analysis Regression Tables for the United States

### Appendix 2.1: Value Weighted T20V and B20V Portfolios (U.S.)

**Table 24 :** Fama-French 5-factor loading on value-weighted T20V, B20V and L-S portfolios based on S&P 500 constituents

	2010 - 2023			2010 - 2016			2017 - 2023		
	T20V	B20V	L-S	T20V	B20V	L-S	T20V	B20V	L-S
RMt-Rft	0.9960*** (0.0320)	1.0701*** (0.0235)	-0.0741* (0.0417)	1.0160*** (0.0228)	1.0909*** (0.0428)	-0.0749 (0.0476)	1.0007*** (0.0472)	1.0814*** (0.0269)	-0.0807 (0.0577)
SMB	-0.0845** (0.0361)	-0.1378*** (0.0488)	0.0532 (0.0662)	-0.1321*** (0.0437)	-0.0885 (0.0657)	-0.0436 (0.0871)	-0.0555 (0.0646)	-0.1662** (0.0701)	0.1106 (0.1140)
HML	0.0125 (0.0548)	-0.0854** (0.0331)	0.0980 (0.0784)	-0.0753 (0.0646)	-0.0352 (0.0712)	-0.0400 (0.1070)	0.0341 (0.0768)	-0.0963** (0.0418)	0.1304 (0.1091)
RMW	0.0672 (0.0800)	0.1106* (0.0583)	-0.0433 (0.1228)	0.1139* (0.0676)	0.2398** (0.1149)	-0.1259 (0.1432)	0.0353 (0.1048)	0.0792 (0.0719)	-0.0439 (0.1598)
CMA	0.1682** (0.0756)	-0.1601* (0.0826)	0.3283** (0.1301)	0.0891 (0.0891)	-0.5076*** (0.1096)	0.5966*** (0.1685)	0.1967* (0.1083)	-0.0669 (0.1032)	0.2636 (0.1798)
Alpha	-0.0020** (0.0009)	0.0012 (0.0009)	-0.0032** (0.0015)	-0.0020* (0.0010)	0.0023 (0.0014)	-0.0043** (0.0021)	-0.0018 (0.0014)	0.0001 (0.0011)	-0.0020 (0.0021)
Observations	168	168	168	84	84	84	84	84	84
Adj. R <sup>2</sup>	0.9379	0.9424	0.1939	0.9564	0.9264	0.1234	0.9294	0.9569	0.2280

Note: This table illustrates the Fama-French 5-factor loading on the value-weighted T20V, B20V, and L-S Portfolios, constructed using constituents from the S&P 500 as the underlying basis. The data is categorised into three distinct time periods: 2010-2023, 2010-2016, and 2017-2023. HAC robust standard errors are presented in parentheses. Levels of statistical significance are denoted by \*\*\* (1%), \*\* (5%), and \* (10%).

**Table 25 :** Fama-French 5-factor loading on value-weighted T20V, B20V and L-S portfolios based on S&P 500 constituents during the COVID-19 pandemic

	Entire period			Drawdown period			Recovery period		
	T20V	B20V	L-S	T20V	B20V	L-S	T20V	B20V	L-S
RMt-Rft	1.0076*** (0.0174)	0.9624*** (0.0117)	0.0452* (0.0249)	1.0205*** (0.0154)	0.9643*** (0.0185)	0.0561** (0.0273)	0.9742*** (0.0356)	0.9833*** (0.0276)	-0.0091 (0.0616)
SMB	-0.0285 (0.0525)	-0.1254*** (0.0348)	0.0969 (0.0778)	-0.0063 (0.0844)	-0.1181** (0.0436)	0.1118 (0.1146)	0.0221 (0.0472)	-0.1925*** (0.0491)	0.2146*** (0.0811)
HML	0.1278*** (0.0452)	0.0055 (0.0227)	0.1222** (0.0616)	0.1162 (0.0798)	-0.0343 (0.0318)	0.1505 (0.0971)	0.1205*** (0.0402)	0.0366 (0.0277)	0.0839 (0.0592)
RMW	-0.2174*** (0.0682)	0.3268*** (0.0573)	-0.5442*** (0.1110)	-0.2404* (0.1386)	0.3675*** (0.1087)	-0.6079*** (0.2074)	-0.1603*** (0.0584)	0.2628*** (0.0734)	-0.4230*** (0.1192)
CMA	0.5300*** (0.1071)	-0.2496*** (0.0832)	0.7796*** (0.1652)	0.6697*** (0.1398)	-0.3707*** (0.1161)	1.0405*** (0.1838)	0.4130*** (0.1501)	-0.1562 (0.1144)	0.5691** (0.2454)
Alpha	0.0001 (0.0003)	0.0003 (0.0003)	-0.0002 (0.0005)	0.0014* (0.0007)	0.0001 (0.0005)	0.0013 (0.0011)	-0.0003 (0.0004)	0.0004 (0.0004)	-0.0007 (0.0007)
Observations	135	135	135	34	34	34	101	101	101
Adj. R <sup>2</sup>	0.9830	0.9858	0.3674	0.9938	0.9959	0.5309	0.9436	0.9461	0.3076

Note: This table illustrates the Fama-French 5-factor loading on the value-weighted T20V, B20V, and L-S Portfolios, constructed using constituents from the S&P 500 as the underlying basis. The data is categorised into three distinct time periods: “Entire period” 20/02/2020-31/08/2020, “Drawdown period” 20/02/2020 – 07/04/2020, and “Recovery period” 08/04/2020 – 31/08/2020. HAC robust standard errors are presented in parentheses. Levels of statistical significance are denoted by \*\*\* (1%), \*\* (5%), and \* (10%).

## Appendix 2.2: Carhart 4-Factor Regression (U.S.)

**Table 26 :** Carhart 4-factor loading on equal-weighted T20E, B20E and L-S portfolios based on S&P 500 constituents

	2010 - 2023			2010 - 2016			2017 - 2023		
	T20E	B20E	L-S	T20E	B20E	L-S	T20E	B20E	L-S
RMt-Rft	0.9488*** (0.0221)	1.0011*** (0.0220)	-0.0523** (0.0256)	0.9613*** (0.0222)	1.0536*** (0.0389)	-0.0924*** (0.0339)	0.9565*** (0.0327)	0.9640*** (0.0249)	-0.0074 (0.0358)
SMB	0.0240 (0.0394)	0.1616*** (0.0553)	-0.1376*** (0.0432)	0.0147 (0.0419)	0.1313*** (0.0495)	-0.1165* (0.0667)	0.0265 (0.0698)	0.1435* (0.0824)	-0.1170** (0.0554)
HML	0.2157*** (0.0389)	0.1326*** (0.0405)	0.0831** (0.0357)	-0.0085 (0.0373)	-0.0594 (0.0682)	0.0509 (0.0766)	0.2779*** (0.0389)	0.1847*** (0.0341)	0.0933** (0.0402)
WML	-0.0470 (0.0287)	-0.1262*** (0.0295)	0.0792** (0.0331)	-0.0605 (0.0408)	-0.0428 (0.0456)	-0.0177 (0.0637)	-0.0483 (0.0419)	-0.2083*** (0.0374)	0.1600*** (0.0386)
Alpha	0.0009 (0.0009)	-0.0005 (0.0010)	0.0014 (0.0010)	0.0006 (0.0010)	0.0002 (0.0012)	0.0004 (0.0014)	0.0014 (0.0013)	-0.0012 (0.0015)	0.0026** (0.0012)
Observations	168	168	168	84	84	84	84	84	84
Adj. R <sup>2</sup>	0.9486	0.9470	0.2222	0.9614	0.9601	0.1580	0.9502	0.9503	0.3160

Note: This table illustrates the Carhart 4-factor loading on the equal-weighted T20E, B20E, and L-S Portfolios, constructed using constituents from the S&P 500 as the underlying basis. The data is categorised into three distinct time periods: 2010-2023, 2010-2016, and 2017-2023. HAC robust standard errors are presented in parentheses. Levels of statistical significance are denoted by \*\*\* (1%), \*\* (5%), and \* (10%).

**Table 27 :** Carhart 4-factor loading on equal-weighted T20E, B20E and L-S portfolios based on S&P 500 constituents during the COVID-19 pandemic

	Entire period			Drawdown period			Recovery period		
	T20E	B20E	L-S	T20E	B20E	L-S	T20E	B20E	L-S
RMt-Rft	0.9781*** (0.0151)	0.9877*** (0.0193)	-0.0096 (0.0119)	0.9734*** (0.0207)	0.9592*** (0.0337)	0.0143 (0.0224)	0.9431*** (0.0429)	0.9554*** (0.0306)	-0.0123 (0.0303)
SMB	0.0494 (0.0613)	0.1449*** (0.0489)	-0.0956** (0.0394)	0.0135 (0.0925)	0.1069 (0.0808)	-0.0934 (0.0711)	0.1313** (0.0523)	0.1713*** (0.0432)	-0.0400 (0.0481)
HML	0.3245*** (0.0689)	0.2352*** (0.0660)	0.0893** (0.0359)	0.3914*** (0.1341)	0.3811*** (0.1372)	0.0103 (0.0845)	0.2521*** (0.0586)	0.1323*** (0.0432)	0.1198*** (0.0450)
WML	-0.0012 (0.0398)	-0.0892** (0.0394)	0.0879*** (0.0291)	0.0410 (0.0879)	0.1167 (0.1530)	-0.0756 (0.1264)	-0.0371 (0.0518)	-0.1690*** (0.0363)	0.1319*** (0.0393)
Alpha	0.0002 (0.0004)	-0.0001 (0.0004)	0.0003 (0.0004)	0.0014 (0.0013)	-0.0001 (0.0016)	0.0015 (0.0011)	-0.0002 (0.0003)	-0.0002 (0.0003)	-0.0001 (0.0003)
Observations	135	135	135	34	34	34	101	101	101
Adj. R <sup>2</sup>	0.9802	0.9811	0.1092	0.9892	0.9836	-0.0238	0.9539	0.9738	0.1786

Note: This table illustrates the Carhart 4-factor loading on the equal-weighted T20E, B20E, and L-S Portfolios, constructed using constituents from the S&P 500 as the underlying basis. The data is categorised into three distinct time periods: "Entire period" 20/02/2020-31/08/2020, "Drawdown period" 20/02/2020 – 07/04/2020, and "Recovery period" 08/04/2020 – 31/08/2020. HAC robust standard errors are presented in parentheses. Levels of statistical significance are denoted by \*\*\* (1%), \*\* (5%), and \* (10%).

**Table 28 :** Carhart 4-factor loading on ESGE – MKTE and ESGV – MKTV difference portfolios using S&P 500 constituents

	2010 - 2023		2010 - 2016		2017 - 2023	
	ESGE - MKTE	ESGV - MKTV	ESGE - MKTE	ESGV - MKTV	ESGE - MKTE	ESGV - MKTV
RMt-Rft	-0.0067** (0.0031)	-0.0124*** (0.0043)	-0.0126*** (0.0043)	-0.0030 (0.0040)	-0.0009 (0.0031)	-0.0160** (0.0068)
SMB	-0.0181*** (0.0043)	0.0038 (0.0069)	-0.0191*** (0.0072)	-0.0068 (0.0057)	-0.0139*** (0.0045)	0.0132 (0.0125)
HML	0.0067** (0.0032)	0.0227*** (0.0066)	0.0005 (0.0071)	0.0031 (0.0077)	0.0089*** (0.0029)	0.0259*** (0.0082)
WML	0.0049* (0.0029)	0.0004 (0.0046)	-0.0002 (0.0050)	-0.0099** (0.0049)	0.0117*** (0.0034)	0.0053 (0.0075)
Alpha	0.0001 (0.0001)	-0.0002 (0.0002)	-0.0000 (0.0002)	-0.0003 (0.0002)	0.0002** (0.0001)	-0.0000 (0.0002)
Observations	168	168	84	84	84	84
Adj. R <sup>2</sup>	0.2509	0.1651	0.2694	0.0050	0.3102	0.2735

Note: This table illustrates the Carhart 4-factor loading on the difference portfolios ESGE – MKTE (formed by buying ESGE and selling MKTE) and ESGV – MKTV (formed by buying ESGV and selling MKTV). These portfolios are constructed using constituents from the S&P 500 as the underlying basis. The data is categorised into three distinct time periods: 2010-2023, 2010-2016, and 2017-2023. HAC robust standard errors are presented in parentheses. Levels of statistical significance are denoted by \*\*\* (1%), \*\* (5%), and \* (10%).

## Appendix 2.3: Exclusion of Outliers (U.S.)

**Table 29 :** Fama-French 5-factor loading on equal-weighted T20E, B20E and L-S portfolios based on S&P 500 constituents excluding outliers

	2010 - 2023			2010 - 2016			2017 - 2023		
	T20E	B20E	L-S	T20E	B20E	L-S	T20E	B20E	L-S
RMt-Rft	0.9682*** (0.0253)	1.0372*** (0.0220)	-0.0690*** (0.0222)	0.9822*** (0.0286)	1.0522*** (0.0381)	-0.0700** (0.0328)	0.9796*** (0.0358)	1.0327*** (0.0294)	-0.0530* (0.0298)
SMB	0.0802** (0.0385)	0.2079*** (0.0629)	-0.1276*** (0.0484)	0.0454 (0.0432)	0.1056* (0.0541)	-0.0603 (0.0620)	0.0774 (0.0648)	0.2684** (0.1068)	-0.1910** (0.0787)
HML	0.1567*** (0.0568)	0.1248** (0.0539)	0.0319 (0.0441)	-0.0405 (0.0502)	-0.0666 (0.0722)	0.0262 (0.0896)	0.2238*** (0.0701)	0.1516** (0.0695)	0.0722 (0.0613)
RMW	0.1071 (0.0718)	0.0680 (0.0545)	0.0391 (0.0560)	0.1265 (0.0814)	-0.0950 (0.0824)	0.2215** (0.0947)	0.0479 (0.0817)	0.1105 (0.0731)	-0.0626 (0.0741)
CMA	0.1188 (0.0835)	0.0349 (0.0846)	0.0839 (0.0685)	0.1422 (0.1025)	0.0478 (0.0848)	0.0944 (0.1436)	0.1070 (0.1079)	0.0528 (0.1155)	0.0542 (0.0772)
Alpha	0.0003 (0.0009)	-0.0013 (0.0011)	0.0016 (0.0010)	-0.0000 (0.0010)	0.0001 (0.0011)	-0.0002 (0.0013)	0.0009 (0.0013)	-0.0025 (0.0018)	0.0034** (0.0014)
Observations	168	168	168	84	84	84	84	84	84
Adj. R <sup>2</sup>	0.9498	0.9414	0.1922	0.9626	0.9596	0.2125	0.9499	0.9370	0.1860

Note: This table illustrates the Fama-French 5-factor loading on the equal-weighted T20E, B20E, and L-S Portfolios, constructed using constituents from the S&P 500 excluding outliers as the underlying basis. The data is categorised into three distinct time periods: 2010-2023, 2010-2016, and 2017-2023. HAC robust standard errors are presented in parentheses. Levels of statistical significance are denoted by \*\*\* (1%), \*\* (5%), and \* (10%).

**Table 30 : Fama-French 5-factor loading on equal-weighted T20E, B20E and L-S portfolios based on S&P 500 constituents during the COVID-19 pandemic excluding outliers**

	Entire period			Drawdown period			Recovery period		
	T20E	B20E	L-S	T20E	B20E	L-S	T20E	B20E	L-S
RMt-Rft	0.9981*** (0.0164)	0.9751*** (0.0208)	0.0230* (0.0123)	0.9928*** (0.0157)	0.9657*** (0.0180)	0.0271 (0.0185)	0.9885*** (0.0380)	0.9876*** (0.0255)	0.0009 (0.0268)
SMB	0.0909 (0.0609)	0.1784*** (0.0595)	-0.0875** (0.0362)	0.0755 (0.1066)	0.1428* (0.0762)	-0.0674 (0.0938)	0.1365*** (0.0515)	0.2275*** (0.0552)	-0.0910* (0.0506)
HML	0.2903*** (0.0470)	0.2849*** (0.0443)	0.0054 (0.0323)	0.3430*** (0.1100)	0.3050** (0.1304)	0.0381 (0.0815)	0.2581*** (0.0396)	0.2674*** (0.0378)	-0.0093 (0.0333)
RMW	-0.1618* (0.0843)	0.2112** (0.0844)	-0.3731*** (0.0579)	-0.1574 (0.2168)	0.3036 (0.2008)	-0.4610*** (0.1184)	-0.1231* (0.0664)	0.1842*** (0.0642)	-0.3073*** (0.0692)
CMA	0.3611*** (0.1180)	-0.0216 (0.1086)	0.3827*** (0.0886)	0.3155 (0.2167)	-0.0460 (0.2036)	0.3615* (0.1964)	0.3742*** (0.1399)	-0.0354 (0.1200)	0.4095*** (0.0933)
Alpha	0.0002 (0.0004)	-0.0000 (0.0004)	0.0003 (0.0003)	0.0014 (0.0014)	0.0002 (0.0017)	0.0013 (0.0010)	-0.0002 (0.0003)	-0.0002 (0.0003)	0.0000 (0.0003)
Observations	135	135	135	34	34	34	101	101	101
Adj. R <sup>2</sup>	0.9825	0.9811	0.3113	0.9898	0.9839	0.2516	0.9586	0.9688	0.2994

Note: This table illustrates the Fama-French 5-factor loading on the equal-weighted T20E, B20E, and L-S Portfolios, constructed using constituents from the S&P 500, excluding outliers as the underlying basis. The data is categorised into three distinct time periods: “Entire period” 20/02/2020-31/08/2020, “Drawdown period” 20/02/2020 – 07/04/2020, and “Recovery period” 08/04/2020 – 31/08/2020. HAC robust standard errors are presented in parentheses. Levels of statistical significance are denoted by \*\*\* (1%), \*\* (5%), and \* (10%).

**Table 31 : Fama-French 5-factor loading on ESGE – MKTE and ESGV – MKTV difference portfolios using S&P 500 constituents excluding outliers**

	2010 - 2023		2010 - 2016		2017 - 2023	
	ESGE - MKTE	ESGV - MKTV	ESGE - MKTE	ESGV - MKTV	ESGE - MKTE	ESGV - MKTV
RMt-Rft	-0.0079*** (0.0027)	-0.0100** (0.0045)	-0.0100** (0.0042)	-0.0006 (0.0041)	-0.0046 (0.0028)	-0.0151** (0.0063)
SMB	-0.0141*** (0.0049)	0.0083 (0.0065)	-0.0113 (0.0071)	-0.0045 (0.0066)	-0.0187*** (0.0063)	0.0169 (0.0117)
HML	0.0016 (0.0048)	0.0057 (0.0081)	-0.0058 (0.0103)	-0.0135 (0.0110)	0.0077 (0.0054)	0.0113 (0.0108)
RMW	0.0095 (0.0059)	0.0054 (0.0120)	0.0298*** (0.0109)	0.0082 (0.0104)	-0.0044 (0.0051)	0.0040 (0.0160)
CMA	0.0075 (0.0074)	0.0381*** (0.0139)	0.0199 (0.0181)	0.0533*** (0.0163)	0.0026 (0.0074)	0.0309 (0.0190)
Alpha	0.0001 (0.0001)	-0.0002* (0.0001)	-0.0001 (0.0002)	-0.0004** (0.0002)	0.0003*** (0.0001)	-0.0001 (0.0002)
Observations	168	168	84	84	84	84
Adj. R <sup>2</sup>	0.2514	0.2323	0.3599	0.1128	0.2106	0.3054

Note: This table illustrates the Fama-French 5-factor loading on the difference portfolios ESGE – MKTE (formed by buying ESGE and selling MKTE) and ESGV – MKTV (formed by buying ESGV and selling MKTV). These portfolios are constructed using constituents from the S&P 500, excluding outliers as the underlying basis. The data is categorised into three distinct time periods: 2010-2023, 2010-2016, and 2017-2023. HAC robust standard errors are presented in parentheses. Levels of statistical significance are denoted by \*\*\* (1%), \*\* (5%), and \* (10%).

## Appendix 3: Robustness Analysis Regression Tables for Europe

### Appendix 3.1: Value Weighted T20V and B20V Portfolios (Europe)

**Table 32 :** Fama-French 5-factor loading on value-weighted T20V, B20V and L-S portfolios based on STOXX Europe 600 constituents

	2010 - 2023			2010 - 2016			2017 - 2023		
	T20V	B20V	L-S	T20V	B20V	L-S	T20V	B20V	L-S
RMt-Rft	1.0379*** (0.0227)	1.0757*** (0.0183)	-0.0378 (0.0266)	1.0544*** (0.0347)	1.1026*** (0.0284)	-0.0482 (0.0380)	0.9910*** (0.0289)	1.0991*** (0.0292)	-0.1081*** (0.0309)
SMB	-0.2693*** (0.0476)	0.1890*** (0.0494)	-0.4582*** (0.0714)	-0.2825*** (0.0422)	0.1949*** (0.0557)	-0.4774*** (0.0681)	-0.1305* (0.0746)	0.0926 (0.0760)	-0.2232* (0.1132)
HML	-0.0677 (0.0671)	0.2453*** (0.0837)	-0.3130*** (0.1076)	0.1247 (0.0926)	-0.1017 (0.1417)	0.2263 (0.1527)	-0.1077 (0.0761)	0.3261*** (0.0791)	-0.4339*** (0.0921)
RMW	-0.0201 (0.0835)	0.3376*** (0.0918)	-0.3577*** (0.1315)	0.1700 (0.1074)	0.0306 (0.1537)	0.1393 (0.1901)	0.0338 (0.0956)	0.3246** (0.1370)	-0.2908* (0.1735)
CMA	0.1260 (0.0938)	-0.0057 (0.1059)	0.1317 (0.1467)	0.1994* (0.1116)	-0.1779 (0.1247)	0.3774*** (0.1385)	0.1164 (0.1246)	0.0403 (0.1112)	0.0762 (0.1509)
Alpha	0.0005 (0.0007)	-0.0013 (0.0009)	0.0018 (0.0013)	-0.0002 (0.0011)	0.0010 (0.0009)	-0.0012 (0.0016)	0.0011 (0.0008)	-0.0032*** (0.0009)	0.0043*** (0.0013)
Observations	168	168	168	84	84	84	84	84	84
Adj. R <sup>2</sup>	0.9717	0.9591	0.2571	0.9754	0.9715	0.4558	0.9747	0.9621	0.4341

Note: This table illustrates the Fama-French 5-factor loading on the value-weighted T20V, B20V, and L-S Portfolios, constructed using constituents from the STOXX Europe 600 as the underlying basis. The data is categorised into three distinct time periods: 2010-2023, 2010-2016, and 2017-2023. HAC robust standard errors are presented in parentheses. Levels of statistical significance are denoted by \*\*\* (1%), \*\* (5%), and \* (10%).

**Table 33 :** Fama-French 5-factor loading on value-weighted T20V, B20V and L-S portfolios based on S&P 500 constituents during the COVID-19 pandemic

	Entire period			Drawdown period			Recovery period		
	T20V	B20V	L-S	T20V	B20V	L-S	T20V	B20V	L-S
RMt-Rft	1.0129*** (0.0170)	1.1530*** (0.0283)	-0.1401*** (0.0259)	1.0520*** (0.0271)	1.1399*** (0.0525)	-0.0880** (0.0377)	1.0145*** (0.0457)	1.2021*** (0.0549)	-0.1876*** (0.0359)
SMB	-0.1886*** (0.0518)	0.0279 (0.0806)	-0.2166*** (0.0700)	-0.2283*** (0.0636)	0.0259 (0.1188)	-0.2542** (0.1124)	-0.0670 (0.0977)	0.1586 (0.1400)	-0.2257** (0.1021)
HML	-0.1621** (0.0721)	0.5110*** (0.0962)	-0.6731*** (0.1003)	-0.2885** (0.1113)	0.5539** (0.2022)	-0.8424*** (0.1835)	-0.1400 (0.1053)	0.5122*** (0.1163)	-0.6522*** (0.0969)
RMW	0.0278 (0.1291)	0.3192* (0.1854)	-0.2914 (0.1780)	-0.0631 (0.3198)	0.5253 (0.3927)	-0.5885 (0.4135)	-0.0063 (0.1313)	0.2051 (0.1798)	-0.2114 (0.1536)
CMA	0.0748 (0.1242)	0.3794** (0.1783)	-0.3046 (0.1848)	-0.1224 (0.1740)	0.5686** (0.2703)	-0.6910** (0.2547)	0.1457 (0.1779)	0.2891 (0.2281)	-0.1434 (0.1616)
Alpha	0.0002 (0.0002)	0.0009** (0.0004)	-0.0007* (0.0004)	0.0002 (0.0007)	0.0021 (0.0013)	-0.0018 (0.0012)	0.0001 (0.0003)	0.0003 (0.0004)	-0.0003 (0.0004)
Observations	138	138	138	34	34	34	104	104	104
Adj. R <sup>2</sup>	0.9701	0.9706	0.8070	0.9846	0.9803	0.8289	0.9380	0.9487	0.7946

Note: This table illustrates the Fama-French 5-factor loading on the value-weighted T20V, B20V, and L-S Portfolios, constructed using constituents from the STOXX Europe 600 as the underlying basis. The data is categorised into three distinct time periods: “Entire period” 20/02/2020-31/08/2020, “Drawdown period” 20/02/2020 – 07/04/2020, and “Recovery period” 08/04/2020 – 31/08/2020. HAC robust standard errors are presented in parentheses. Levels of statistical significance are denoted by \*\*\* (1%), \*\* (5%), and \* (10%).

## Appendix 3.2: Carhart 4-Factor Regression (Europe)

**Table 34 :** Carhart 4-factor loading on equal-weighted T20E, B20E and L-S portfolios based on STOXX Europe 600 constituents

	2010 - 2023			2010 - 2016			2017 - 2023		
	T20E	B20E	L-S	T20E	B20E	L-S	T20E	B20E	L-S
RMt-Rft	1.0829*** (0.0189)	1.1256*** (0.0265)	-0.0426* (0.0232)	1.0610*** (0.0368)	1.0979*** (0.0327)	-0.0369 (0.0334)	1.0620*** (0.0172)	1.1386*** (0.0285)	-0.0766*** (0.0277)
SMB	0.0430 (0.0443)	0.4948*** (0.0467)	-0.4518*** (0.0543)	-0.0759 (0.0462)	0.4318*** (0.0625)	-0.5077*** (0.0678)	0.2033*** (0.0590)	0.5148*** (0.0755)	-0.3115*** (0.0733)
HML	0.0999*** (0.0292)	-0.1087*** (0.0375)	0.2085*** (0.0371)	0.1427** (0.0713)	-0.0955 (0.0645)	0.2381*** (0.0790)	0.0888*** (0.0273)	-0.0912** (0.0421)	0.1800*** (0.0436)
WML	-0.0924*** (0.0294)	-0.1336*** (0.0419)	0.0411 (0.0432)	-0.0579 (0.0375)	-0.0562 (0.0668)	-0.0017 (0.0600)	-0.1349*** (0.0336)	-0.1902*** (0.0473)	0.0553 (0.0646)
Alpha	0.0008 (0.0006)	0.0004 (0.0010)	0.0004 (0.0010)	0.0018** (0.0007)	0.0007 (0.0013)	0.0011 (0.0014)	0.0005 (0.0008)	-0.0001 (0.0013)	0.0006 (0.0014)
Observations	168	168	168	84	84	84	84	84	84
Adj. R <sup>2</sup>	0.9799	0.9681	0.3755	0.9782	0.9628	0.4058	0.9841	0.9741	0.3637

Note: This table illustrates the Carhart 4-factor loading on the equal-weighted T20E, B20E, and L-S Portfolios, constructed using constituents from the STOXX Europe 600 as the underlying basis. The data is categorised into three distinct time periods: 2010-2023, 2010-2016, and 2017-2023. HAC robust standard errors are presented in parentheses. Levels of statistical significance are denoted by \*\*\* (1%), \*\* (5%), and \* (10%).

**Table 35 :** Carhart 4-factor loading on equal-weighted T20E, B20E and L-S portfolios based on STOXX Europe 600 constituents during the COVID-19 pandemic

	Entire period			Drawdown period			Recovery period		
	T20E	B20E	L-S	T20E	B20E	L-S	T20E	B20E	L-S
RMt-Rft	1.1530*** (0.0308)	1.2149*** (0.0312)	-0.0619** (0.0238)	1.2465*** (0.0507)	1.2753*** (0.0596)	-0.0288 (0.0479)	1.0522*** (0.0699)	1.1337*** (0.0573)	-0.0814** (0.0381)
SMB	0.2653*** (0.0728)	0.5263*** (0.0801)	-0.2610*** (0.0524)	0.3626*** (0.1178)	0.6054*** (0.1169)	-0.2428** (0.0959)	0.2024 (0.1225)	0.4724*** (0.1393)	-0.2700*** (0.0875)
HML	-0.0307 (0.1003)	-0.0778 (0.1105)	0.0470 (0.0816)	-0.3544* (0.1754)	-0.2603 (0.2094)	-0.0941 (0.1890)	-0.0351 (0.1462)	-0.1627 (0.1511)	0.1276 (0.0859)
WML	-0.2476*** (0.0672)	-0.2938*** (0.0659)	0.0462 (0.0597)	-0.4156** (0.1555)	-0.5578*** (0.1101)	0.1422 (0.1709)	-0.3018*** (0.0938)	-0.3595*** (0.1000)	0.0577 (0.0664)
Alpha	-0.0001 (0.0003)	0.0003 (0.0004)	-0.0003 (0.0003)	-0.0001 (0.0009)	0.0008 (0.0011)	-0.0009 (0.0009)	0.0001 (0.0004)	0.0002 (0.0004)	-0.0001 (0.0004)
Observations	138	138	138	34	34	34	104	104	104
Adj. R <sup>2</sup>	0.9668	0.9617	0.1091	0.9830	0.9749	0.1430	0.9373	0.9343	0.0792

Note: This table illustrates the Carhart 4-factor loading on the equal-weighted T20E, B20E, and L-S Portfolios, constructed using constituents from the STOXX Europe 600 as the underlying basis. The data is categorised into three distinct time periods: "Entire period" 20/02/2020-31/08/2020, "Drawdown period" 20/02/2020 – 07/04/2020, and "Recovery period" 08/04/2020 – 31/08/2020. HAC robust standard errors are presented in parentheses. Levels of statistical significance are denoted by \*\*\* (1%), \*\* (5%), and \* (10%).

**Table 36 :** Carhart 4-factor loading on ESGE – MKTE and ESGV – MKTV difference portfolios using STOXX Europe 600 constituents

	2010 - 2023		2010 - 2016		2017 - 2023	
	ESGE - MKTE	ESGV - MKTV	ESGE - MKTE	ESGV - MKTV	ESGE - MKTE	ESGV - MKTV
RMt-Rft	-0.0015 (0.0022)	0.0017 (0.0030)	-0.0029 (0.0036)	-0.0002 (0.0049)	-0.0050** (0.0022)	-0.0049 (0.0037)
SMB	-0.0500*** (0.0055)	-0.0355*** (0.0067)	-0.0606*** (0.0078)	-0.0374*** (0.0062)	-0.0299*** (0.0057)	-0.0194* (0.0110)
HML	0.0182*** (0.0035)	-0.0104* (0.0061)	0.0250*** (0.0088)	0.0209* (0.0119)	0.0144*** (0.0036)	-0.0271*** (0.0043)
WML	0.0035 (0.0043)	0.0086 (0.0070)	0.0005 (0.0077)	0.0034 (0.0076)	0.0041 (0.0045)	0.0120** (0.0055)
Alpha	0.0001 (0.0001)	-0.0000 (0.0002)	0.0001 (0.0002)	-0.0000 (0.0002)	0.0001 (0.0001)	0.0002 (0.0002)
Observations	168	168	84	84	84	84
Adj. R <sup>2</sup>	0.3676	0.1768	0.4039	0.2247	0.3470	0.5055

Note: This table illustrates the Carhart 4-factor loading on the difference portfolios ESGE – MKTE (formed by buying ESGE and selling MKTE) and ESGV – MKTV (formed by buying ESGV and selling MKTV). These portfolios are constructed using constituents from the STOXX Europe 600 as the underlying basis. The data is categorised into three distinct time periods: 2010-2023, 2010-2016, and 2017-2023. HAC robust standard errors are presented in parentheses. Levels of statistical significance are denoted by \*\*\* (1%), \*\* (5%), and \* (10%).

### Appendix 3.3: Exclusion of Outliers (Europe)

**Table 37 :** Fama-French 5-factor loading on equal-weighted T20E, B20E and L-S portfolios based on STOXX Europe 600 constituents excluding outliers

	2010 - 2023			2010 - 2016			2017 - 2023		
	T20E	B20E	L-S	T20E	B20E	L-S	T20E	B20E	L-S
RMt-Rft	1.1032*** (0.0189)	1.1309*** (0.0218)	-0.0277 (0.0232)	1.0729*** (0.0339)	1.0780*** (0.0318)	-0.0052 (0.0374)	1.0960*** (0.0244)	1.1679*** (0.0263)	-0.0719** (0.0299)
SMB	0.0420 (0.0464)	0.4661*** (0.0437)	-0.4241*** (0.0543)	-0.0611 (0.0509)	0.3982*** (0.0583)	-0.4592*** (0.0658)	0.1522** (0.0663)	0.4034*** (0.0699)	-0.2512*** (0.0727)
HML	0.2220*** (0.0546)	0.1138* (0.0598)	0.1082 (0.0743)	0.2407*** (0.0680)	-0.0358 (0.1378)	0.2765** (0.1309)	0.2333*** (0.0762)	0.1780** (0.0748)	0.0553 (0.1019)
RMW	0.1822*** (0.0656)	0.1714* (0.0921)	0.0108 (0.0979)	0.1308 (0.0864)	-0.0988 (0.1581)	0.2296 (0.1599)	0.2116** (0.0980)	0.2154** (0.1076)	-0.0039 (0.1285)
CMA	-0.0471 (0.0791)	-0.2545** (0.1028)	0.2074* (0.1184)	-0.0045 (0.1260)	-0.2584* (0.1431)	0.2539 (0.1558)	-0.0856 (0.1114)	-0.3068** (0.1434)	0.2212 (0.1689)
Alpha	-0.0006 (0.0006)	-0.0011 (0.0009)	0.0005 (0.0008)	0.0008 (0.0009)	0.0012 (0.0009)	-0.0003 (0.0013)	-0.0013 (0.0008)	-0.0025** (0.0012)	0.0012 (0.0010)
Observations	168	168	168	84	84	84	84	84	84
Adj. R <sup>2</sup>	0.9786	0.9675	0.3712	0.9772	0.9650	0.4454	0.9811	0.9716	0.3389

Note: This table illustrates the Fama-French 5-factor loading on the equal-weighted T20E, B20E, and L-S Portfolios, constructed using constituents from the STOXX Europe 600 excluding outliers as the underlying basis. The data is categorised into three distinct time periods: 2010-2023, 2010-2016, and 2017-2023. HAC robust standard errors are presented in parentheses. Levels of statistical significance are denoted by \*\*\* (1%), \*\* (5%), and \* (10%).

**Table 38 : Fama-French 5-factor loading on equal-weighted T20E, B20E and L-S portfolios based on STOXX Europe 600 constituents excluding outliers during the COVID-19 pandemic**

	Entire period			Drawdown period			Recovery period		
	T20E	B20E	L-S	T20E	B20E	L-S	T20E	B20E	L-S
RMt-Rft	1.1269*** (0.0322)	1.1854*** (0.0357)	-0.0586** (0.0252)	1.1857*** (0.0449)	1.2081*** (0.0718)	-0.0224 (0.0458)	1.1608*** (0.0695)	1.2611*** (0.0684)	-0.1003*** (0.0341)
SMB	0.1956** (0.0788)	0.4122*** (0.0752)	-0.2166*** (0.0635)	0.2555*** (0.0853)	0.4294*** (0.1291)	-0.1740 (0.1045)	0.2928** (0.1470)	0.5718*** (0.1766)	-0.2790*** (0.0967)
HML	0.2463** (0.1054)	0.2558 (0.1587)	-0.0095 (0.1310)	-0.0513 (0.1598)	0.0908 (0.3012)	-0.1421 (0.2351)	0.3715*** (0.1245)	0.3398** (0.1504)	0.0318 (0.1112)
RMW	0.1283 (0.1533)	-0.0214 (0.2798)	0.1496 (0.2989)	-0.0647 (0.2907)	-0.4504 (0.6099)	0.3857 (0.6888)	0.1870 (0.1865)	0.1683 (0.2464)	0.0187 (0.2110)
CMA	0.2123 (0.1688)	0.0738 (0.2159)	0.1385 (0.1556)	0.1329 (0.2128)	-0.0245 (0.2785)	0.1574 (0.1968)	0.1197 (0.2043)	0.0330 (0.2697)	0.0866 (0.1938)
Alpha	0.0004 (0.0004)	0.0008* (0.0005)	-0.0003 (0.0004)	0.0006 (0.0012)	0.0016 (0.0015)	-0.0010 (0.0010)	0.0001 (0.0004)	0.0001 (0.0004)	-0.0000 (0.0004)
Observations	138	138	138	34	34	34	104	104	104
Adj. R <sup>2</sup>	0.9620	0.9548	0.1015	0.9780	0.9680	0.1069	0.9311	0.9241	0.0757

Note: This table illustrates the Fama-French 5-factor loading on the equal-weighted T20E, B20E, and L-S Portfolios, constructed using constituents from the STOXX Europe 600, excluding outliers as the underlying basis. The data is categorised into three distinct time periods: “Entire period” 20/02/2020-31/08/2020, “Drawdown period” 20/02/2020 – 07/04/2020, and “Recovery period” 08/04/2020 – 31/08/2020. HAC robust standard errors are presented in parentheses. Levels of statistical significance are denoted by \*\*\* (1%), \*\* (5%), and \* (10%).

**Table 39 : Fama-French 5-factor loading on ESGE – MKTE and ESGV – MKTV difference portfolios using STOXX Europe 600 constituents excluding outliers**

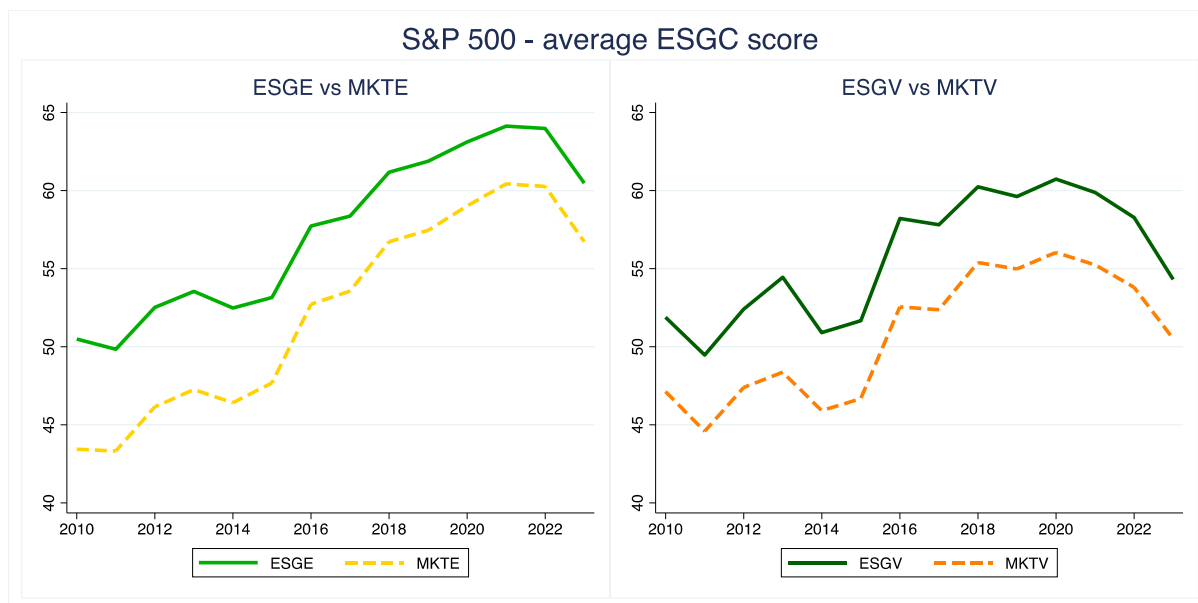
	2010 - 2023		2010 - 2016		2017 - 2023	
	ESGE - MKTE	ESGV - MKTV	ESGE - MKTE	ESGV - MKTV	ESGE - MKTE	ESGV - MKTV
RMt-Rft	0.0000 (0.0020)	0.0010 (0.0030)	0.0015 (0.0039)	0.0039 (0.0043)	-0.0050** (0.0019)	-0.0076*** (0.0027)
SMB	-0.0465*** (0.0055)	-0.0347*** (0.0068)	-0.0535*** (0.0074)	-0.0301*** (0.0067)	-0.0248*** (0.0062)	-0.0158* (0.0084)
HML	0.0120* (0.0072)	-0.0286*** (0.0109)	0.0321* (0.0172)	0.0269 (0.0168)	0.0077 (0.0076)	-0.0407*** (0.0081)
RMW	0.0064 (0.0101)	-0.0239 (0.0145)	0.0348 (0.0219)	0.0368 (0.0223)	0.0002 (0.0108)	-0.0228 (0.0141)
CMA	0.0196 (0.0124)	0.0140 (0.0138)	0.0342** (0.0170)	0.0367** (0.0143)	0.0123 (0.0130)	0.0055 (0.0127)
Alpha	0.0001 (0.0001)	0.0001 (0.0001)	-0.0001 (0.0002)	-0.0002 (0.0002)	0.0002** (0.0001)	0.0004*** (0.0001)
Observations	168	168	84	84	84	84
Adj. R <sup>2</sup>	0.3773	0.1748	0.4731	0.3124	0.3288	0.4762

Note: This table illustrates the Fama-French 5-factor loading on the difference portfolios ESGE – MKTE (formed by buying ESGE and selling MKTE) and ESGV – MKTV (formed by buying ESGV and selling MKTV). These portfolios are constructed using constituents from the STOXX Europe 600, excluding outliers as the underlying basis. The data is categorised into three distinct time periods: 2010-2023, 2010-2016, and 2017-2023. HAC robust standard errors are presented in parentheses. Levels of statistical significance are denoted by \*\*\* (1%), \*\* (5%), and \* (10%).

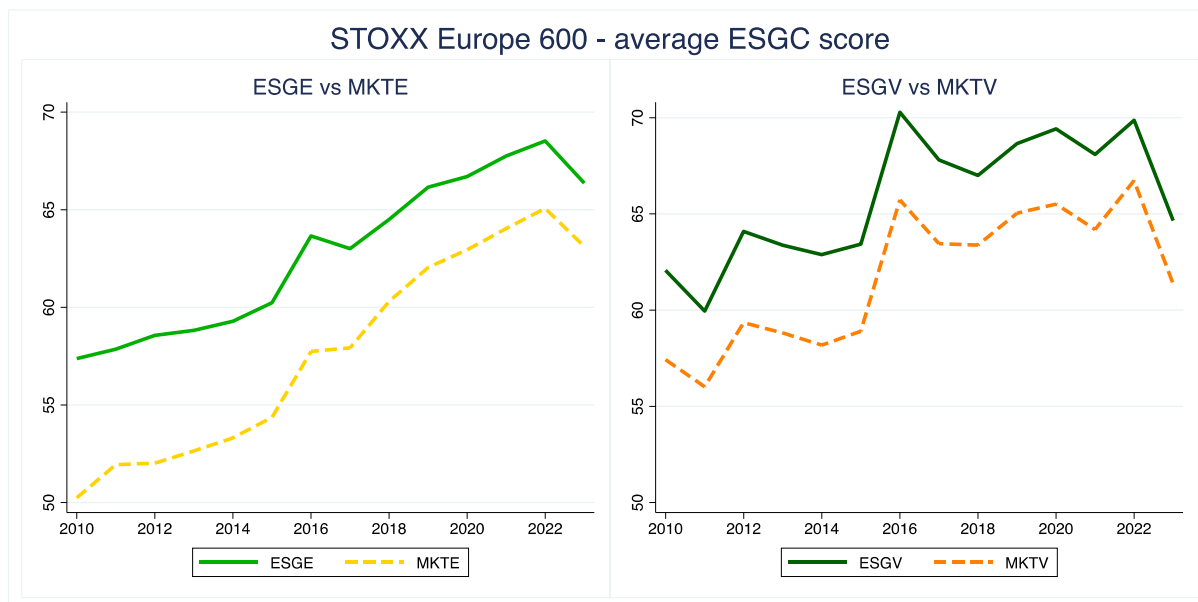
## Appendix 4: Impact of ESGC Weighting on Sustainability

Note: These charts illustrate how integrating ESGC weighting into conventional equal-weighted or value-weighted portfolios affected the weighted average ESGC score of the portfolio. They display the weighted average ESGC score for each of the four distinct asset allocation strategies: ESGE, MKTE, ESGV, and MKTV.

**Figure 9 :** Annual evolution of weighted average ESGC scores in S&P 500 derived portfolios



**Figure 10 :** Annual evolution of weighted average ESGC scores in STOXX Europe 600 derived portfolios



## Appendix 5: Annual Turnover Rates

The annual turnover rate indicates the percentage of assets in each portfolio that are involved in transactions - either buying or selling - during the annual rebalancing process. It is calculated every January 1<sup>st</sup> by comparing the weights of individual assets in the portfolio on January 1<sup>st</sup> of year t to those from December 31<sup>st</sup> of year t-1.

The weight of each asset i at year-end t ( $w_{t,end}$ ) is determined by multiplying its starting weight at year t ( $w_{t,start}$ ) by the cumulative returns over the year, then normalising to ensure the sum of all individual weights equals 100%:

$$w_{ti,end} = \frac{w_{ti,start} \cdot \prod_{m=1}^{12} (1 + R_{tmi})}{\sum_{i=1}^N [w_{ti,start} \cdot \prod_{m=1}^{12} (1 + R_{tmi})]}$$

$R_{tmi}$  represents the return of stock i in month m of year t, and N denotes the total number of stocks in the portfolio at year t.

When comparing the weights of individual assets (hereinafter denoted as i) in the portfolio from January 1<sup>st</sup> of year t (represented as  $w_{t,start}$ ) to those on December 31<sup>st</sup> of year t-1 (represented as  $w_{(t-1),end}$ ) three scenarios may arise:

- 1) When asset i remains in the portfolio for both years, turnover is calculated as the absolute difference between the weights, representing the extent to which the asset needs to be either bought or sold to achieve the required weight in the portfolio:

$$Turnover_{ti} = |w_{ti,start} - w_{(t-1)i,end}|$$

- 2) If asset i exits the portfolio, meaning it is included in the constituent list in year t-1 but not in year t, the asset needs to be sold at the end of the year:

$$Turnover_{ti} = w_{(t-1)i,end}$$

- 3) If asset i enters the portfolio, meaning it was not in the portfolio in year t-1 but is listed in year t, the asset needs to be bought to achieve the required weight in the portfolio:

$$Turnover_{ti} = w_{ti,start}$$

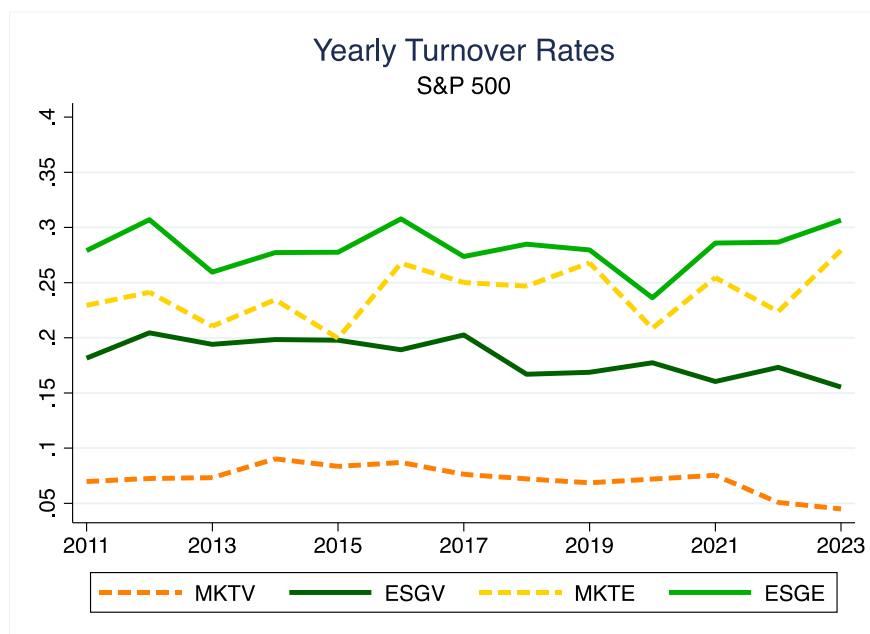
Finally, the turnover rate is calculated by summing all the turnover values for all the new and previous stocks, which are already expressed as percentages since they are derived based on weights.

$$Turnover Rate_t = \sum_i^M Turnover_{ti}$$

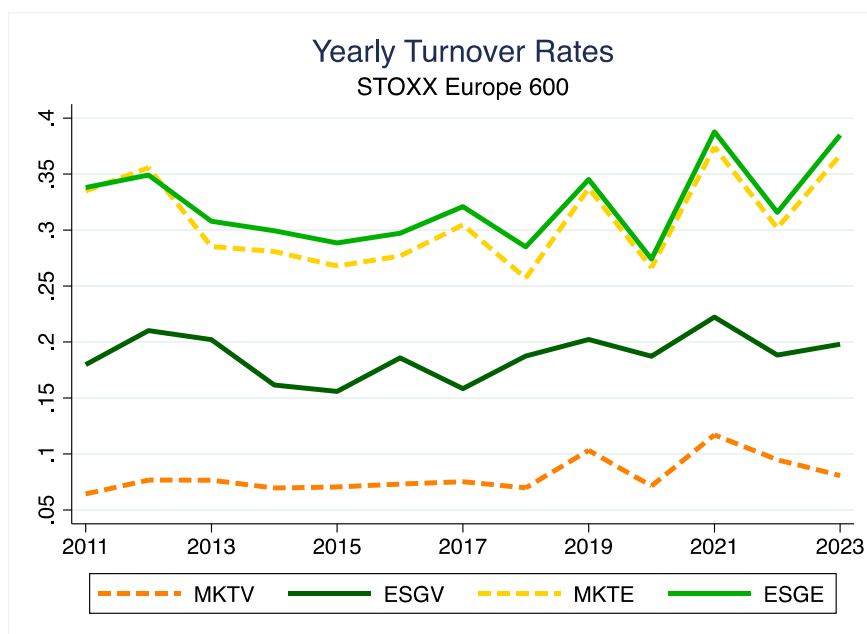
Where M represents the number of unique stocks in the constituent list of year t and t-1.

Figures 11 and 12 illustrate the annual turnover rates for the ESGE, MKTE, ESGV, and MKTV portfolios over the years. Figure 11 presents the annual evolution of turnover rates for S&P 500-derived portfolios and Figure 12 for STOXX Europe 600-derived portfolios.

**Figure 11** : Annual evolution of turnover rates in S&P 500-derived portfolios

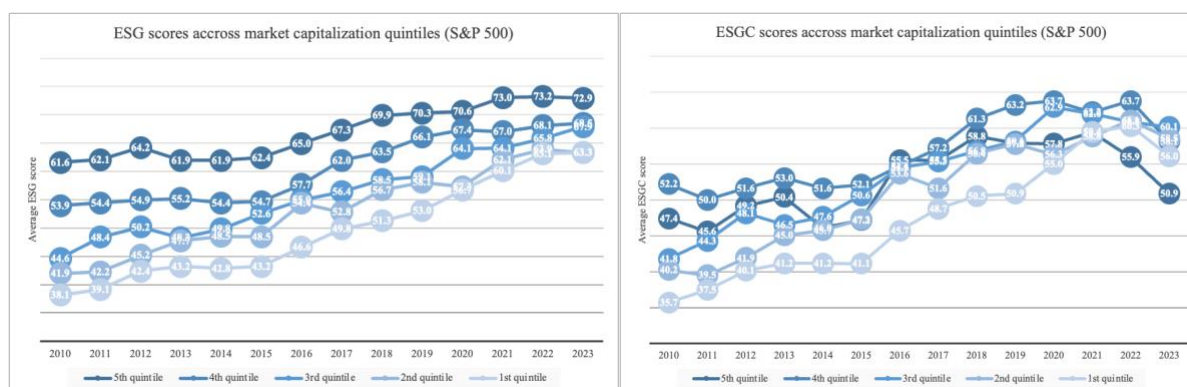


**Figure 12** : Annual evolution of turnover rates STOXX Europe 600-derived portfolios



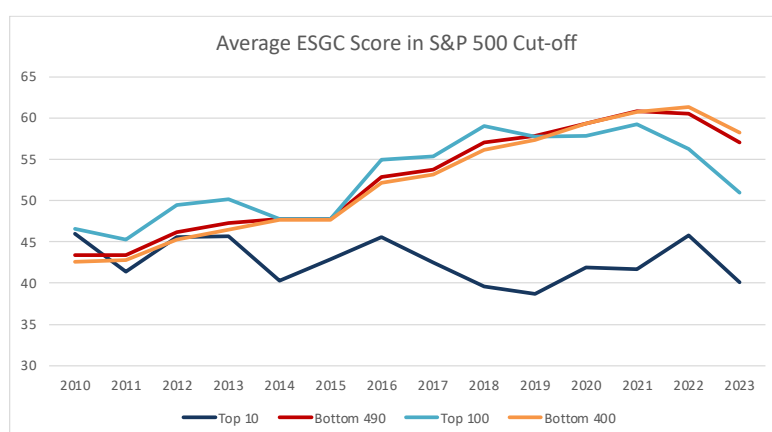
## Appendix 6: ESGC Score Bias – Example of S&P 500

**Figure 13 :** Average ESG and ESGC score over different market capitalisation quintiles in the S&P 500



As discussed in Chapter 2.2, the basic ESG score exhibits a strong upward bias toward large-cap stocks, respectively, a strong downward bias towards small-cap stocks, consistent with the acknowledged limitations of ESG ratings. However, this bias has diminished over time, with only modest discrepancies observed between high and low-performing ESG companies in 2022 and 2023. Compared to the ESG score, the ESGC score presents its own bias with no clear pattern. As evidenced in Figure 13, heightened controversies surrounding the largest companies significantly reduce their ESGC scores, resulting in a bias toward mid-quintile companies, particularly in the 4<sup>th</sup> and 3<sup>rd</sup> quintiles. Interestingly, controversies surrounding companies in the 5<sup>th</sup> quintile have increased over the past three years, resulting in them having the lowest ESGC scores. However, it is not the entire 5<sup>th</sup> quintile (comprising around 100 companies) but primarily the largest companies that face significant controversies, leading to a substantial reduction in their ESGC scores despite initially having the highest simple ESG scores. This can be visualised in Figure 14 below.

**Figure 14 :** Average ESGC score across different market capitalisation cut-offs in the S&P 500 index

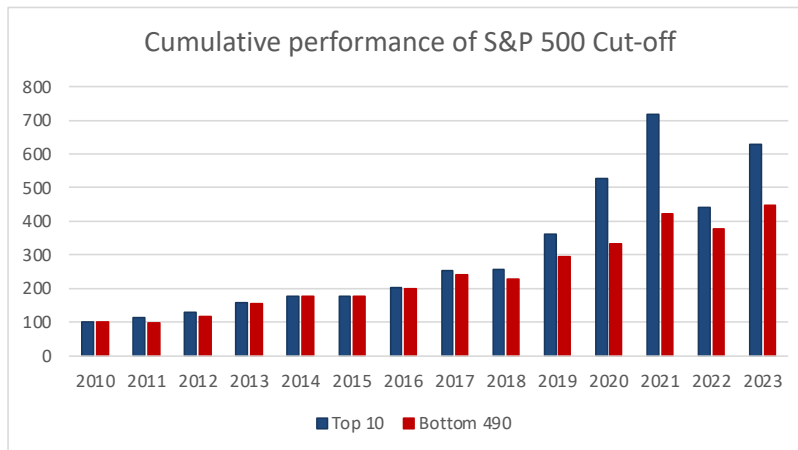


The significant reduction in ESGC scores is primarily observed in the largest of the largest companies, such as the top 10 companies, rather than across the entire 5<sup>th</sup> quintile (top 100 companies by market cap). This discrepancy could pose considerable challenges, especially if the low ESGC scores are not warranted and do not accurately reflect the companies' sustainability

practices. Such discrepancies may lead to biases, particularly when these companies exhibit different return characteristics.

Given that the Top 10 companies represented between 17% and 31% of the total market capitalisation in a market value-weighted index within this study's sample, as opposed to just 2% in an equal-weighted index, it becomes evident that this bias might be negligible for portfolios constructed based on an equal-weighted index, such as T20E, B20E, MKTE, or even ESGE. However, it becomes especially crucial to consider this bias when constructing portfolios based on a value-weighted index, such as T20V, B20V, MKTV, or ESGV.

**Figure 15 :** Cumulative performance of the Top 10 companies by market value in the S&P 500 compared to the remaining 490



A potential concern arises when comparing the return characteristics of the top 10 largest companies in the index with those of the remaining companies. As illustrated in Figure 15, the leading companies in the S&P 500 have outperformed the rest of the index regarding cumulative returns. This suggests that the bias in ESGC scores, which leads to a higher concentration of large-cap companies in the low ESGC portfolio, coincides with the superior performance of these companies. Consequently, this introduces an upward bias in the low ESGC portfolio's performance. Similarly, it results in a downward bias in the ESGV portfolio, as the weight of the largest companies is reduced, ultimately impacting the index's overall performance. While this may not completely explain the negative alphas observed in the S&P 500 value-weighted index, coupled with neutral or positive alphas in the equal-weighted index, it is evident that such a bias could significantly contribute to these findings.

## Appendix 7: ESGC Score Outliers

**Table 40 :** Statistical Outliers based on ESGC score in the STOXX Europe 600 & S&P 500 indices

	Name	ISIN Code	Portfolio Year(s)
STOXX Europe 600	PORSCHE AML.HLDG. (XET) PREF.	DE000PAH0038	2023, 2022, 2021, 2020, 2019, 2018, 2017
	LXI REIT	GB00BYQ46T41	2023
	SMITH (DS)	GB0008220112	2023
	WISE A	GB00BL9YR756	2022
	QT GROUP	FI4000198031	2022
	SOFINA	BE0003717312	2021, 2020, 2019
	TRITAX BIG BOX REIT	GB00BG49KP99	2019
	CD PROJEKT	PLOPTTC00011	2019
	BB BIOTECH N	CH0038389992	2019, 2018, 2017
	PARGESA 'B'	CH0021783391	2018
	AA	GB00BMSKPJ95	2015
	PETROPAVLOVSK	GB0031544546	2012, 2011
S&P 500	TRIPADVISOR 'A'	US8969452015	2018
	(*) TEXAS INSTRUMENTS	US8825081040	2011

Note: This table displays companies identified as statistical outliers based on their ESGC scores within yearly portfolios of constituents from the STOXX Europe 600 and S&P 500 indices. An outlier is defined as a company with an ESGC score falling below the third quartile minus 1.5 times the interquartile range or above the first quartile plus 1.5 times the interquartile range. All companies listed are outliers due to their ESGC scores falling below the minimum threshold, except for the one marked with an asterisk (\*).