Department of Economics and Finance

Course of Econometric Theory

Impact of Sustainable Responsible Investments on Portfolio Performance

Prof. Santucci De Magistris Paolo
SUPERVISOR

Prof. Mazzoni Giancarlo
CO-SUPERVISOR

Giulia Castello
CANDIDATE

Academic Year 2023/2024

Abstract

The rising concern on climate change led to the growth of sustainable responsible investments. Yet, the extant literature yields contrasting results on the risk and return characteristics on those investments. The aim of this thesis is to contribute to the research on SRI through the study of returns of conventional and SRI mutual funds. The applicaton of the Carhart regression model to the rolling performance of a sample of 4000 mutual funds ranked by ESG score and accounting for crisis periods resulted in no positive and statistically significant alpha, meaning no funds are able to gain abnormal returns in normal market conditions. However, lower-ranked funds displayed marginally significant negative alpha over the full period and statistically significant negative alpha during crisis periods, while also the top ranked portfolio showed marginally significant negative alpha in crises, meaning that lower-ranked funds are vulnerable to market downturns, while the top ranked portfolio is not immune to them. The most striking finding concerns the highly statistically significant negative alpha of the regression on a portfolio going long on sustainable funds and short on conventional funds, suggesting poor performance of SRI mutual funds. More in general, the results describe mutual funds returns which are systematic.

Contents

Chapter 1

Introduction

As the world is becoming more and more aware of climate change, the focus on Socially Responsible Investments (SRI) grows. The Global Sustainable Investment Alliance (GSIA), gathering data from the US Sustainable Investment Forum (US SIF), Japan Sustainable Investment Forum (JSIF), the Responsible Investment Association Canada (RIA Canada), the Responsible Investment Association Australasia (RIAA) and the European Fund and Asset Management Association (EFAMA), estimates that in non-US markets, sustainable investment assets under management (AUM) have increased by 20% since 2020 and 30.3 trillion is invested globally in sustainable investing assets (Rayner and Rosanna 2022).

In 2024, the Paris Agreement 1.5°C threshold was exceeded for the first time, with a record of average surface temperatures of 1.6°C above preindustrial levels, with preindustrial period lasting from 1850 to 1900 (Mooney 2025). This does not imply the infringement of the Paris Agreement since the target refer to temperature averages over more than 20 years. However, with the current speed of temperature increase of 0.2°C each decade, it is extremely possible that we will breach the 1.5°C of the Paris Agreement by the 2030s. Moreover, the UN's environment programme recently reported that current government policies globally would lead to 3.1°C (UNEP 2024), while the Intergovernmental Panel on Climate Change has estimated that the temperature rise of the period 2011-2020 is 1.1°C above preindustrial levels (IPCC 2023a). These emissions continue to grow due to unsustainable energy usage, land changes, and consumer trends. Climate change has far-reaching consequences, influencing the weather, oceans, ecosystems, and human populations, with vulnerable communities bearing a disproportionate burden (IPCC 2023a). Despite progress in climate adaptation across sectors, challenges remain, especially with the growing gaps in financial support for adaptation, particularly in developing countries. In terms of mitigation, global efforts are increasing, however current emission reduction policies are still unable to limit global warming to 1.5°C or 2°C (IPCC 2023a). Global warming is expected to continue in the future, with each additional degree of temperature leading to more severe risks and impacts. Deep, rapid, and sustained emission reductions are essential to slow warming and mitigate adverse outcomes. However, reaching net-zero CO2 emissions is critical, as current fossil fuel infrastructure threatens to exceed carbon budgets for keeping warming below 1.5°C (IPCC 2023a). The urgency of integrated climate action is emphasized by the IPCC, as the next decade is crucial for implementing measures that balance both adaptation and mitigation. Accelerated actions will reduce risks to ecosystems and human health, promoting sustainable development. Key enablers for effective climate action include finance, technology, international cooperation, and equitable governance, all of which can enhance the global transition to a low-carbon future (IPCC 2023a).

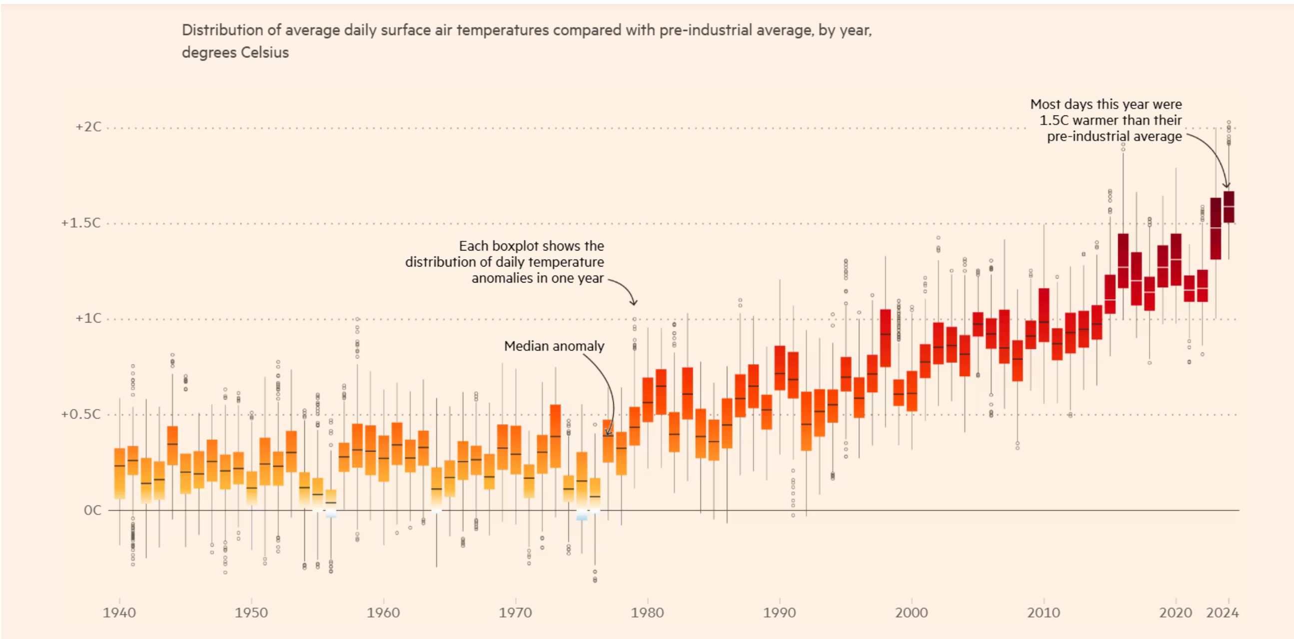

Figure 1.1: Source: ERA5, C3S/ECMWF. Blocks show the interquartile range (IQR). Whiskers represent the daily temperature anomalies within 1.5xIQR from the first or third quartile. Values outside of this are shown as circles.

The IPCC's Sixth Assessment Synthesis Report (IPCC 2023b) highlights the importance of finance in addressing climate change through both mitigation and adaptation efforts. One of the central issues discussed is the significant gap between required and actual financial flows directed toward climate action. Despite increasing commitments to climate finance, current global financial flows remain insufficient to meet the investment needs necessary for limiting warming to 1.5°C or 2°C. A primary concern is the imbalance in funding, with most climate finance directed toward mitigation rather than adaptation. While mitigation efforts receive substantial public and private funding, adaptation measures, particularly in developing nations, continue to be underfunded, exacerbating vulnerabilities to climate-related risks (IPCC 2023a).

The IPCC underscores that financial barriers hinder climate action, particularly in developing economies where access to funding remains constrained by systemic issues such as economic vulnerability, lack of institutional capacity, and regulatory inefficiencies. Public finance plays a crucial role in facilitating adaptation measures, but private investment is necessary to bridge existing financial gaps. Green bonds and sustainable financial products have seen significant growth, yet challenges related to integrity and accessibility in emerging markets persist. The need for an equitable distribution of climate finance is emphasized, particularly as climate-related losses and damages disproportionately affect lower-income nations (IPCC 2023a).

Additionally, the IPCC discusses the need for stronger financial governance and international cooperation. The financial system must integrate climate risk assessments into investment decisions, and regulatory frameworks should be designed to incentivize low-carbon investments. Carbon pricing instruments, though effective in reducing emissions, require higher coverage and price adjustments to drive deeper decarbonization. Furthermore, the removal of fossil fuel subsidies is highlighted as a key measure that could reduce emissions while improving fiscal stability (IPCC 2023a).

The literature on Socially Responsible Investment (SRI) mutual funds, as represented in table A.1 in appendix A, has evolved considerably over the past few decades, focusing on the comparative financial performance of SRI and non-SRI funds. One of the earliest studies in this area by Hamilton et al. (1993) analyzed the performance of 32 SRI mutual funds in the United States between 1981 and 1990, comparing them against a randomly selected sample of 320 conventional funds. Using the Capital Asset Pricing Model (CAPM) and Jensen's alpha against the value-weighted NYSE index, the study found no statistically significant difference in risk-adjusted performance between SRI and conventional funds, suggesting that ethical investing does not necessarily result in underperformance. Schroder (2004) extended the geographical scope of analysis by evaluating European SRI mutual funds from 1990 to 2002. This study employed both CAPM-based measures and the Carhart four-factor model, which includes momentum as an additional explanatory variable. The findings indicated that while SRI funds did not systematically outperform conventional funds, their performance varied significantly based on regional market dynamics and fund-specific screening intensity. Statman (2000) further expanded on the methodology of the research by analyzing US SRI funds alongside traditional benchmarks such as the Domini Social Index (DSI) and the S&P 500. The study determined that SRI funds tended to have higher sectoral concentrations in technology and healthcare, which could introduce factor biases affecting performance comparability. Additionally, the study suggested that investors in SRI funds may accept lower returns in exchange for alignment with ethical preferences, reinforcing the notion of a dual-objective investment strategy. Bauer et al. (2005) undertook a cross-sectional study covering SRI mutual funds in the US, UK, and Germany between 1990 and 2001. Their research applied a time-varying alpha framework and demonstrated that SRI funds exhibited a convergence in performance with conventional funds over time. Notably, the study identified periods where SRI funds outperformed their peers, particularly during economic downturns, suggesting potential resilience to market crises due to sectoral exclusions and ESG-oriented investment strategies. More recent analyses, such as Renneboog et al. (2008b), conducted a global comparison of SRI funds across 17 countries, implementing multifactor models incorporating macroeconomic indicators. The findings indicated substantial regional disparities, with SRI funds in the US and UK performing similarly to conventional funds, whereas those in France and Germany underperformed, possibly due to stricter exclusionary screening. Finally, the work of Gil-Bazo et al. (2010) introduced the aspect of fund management fees into the discussion, demonstrating that SRI funds tend to have higher expense ratios than conventional funds. The study suggested that while these funds may not systematically underperform, their net returns to investors may be lower due to higher management costs associated with ESG research and engagement strategies.

Overall, the literature on SRI mutual fund performance (table A.1) suggests that while SRI funds do not consistently outperform conventional funds, their financial viability is not significantly compromised either. The methodological diversity in these studies—from single-factor CAPM regressions to multifactor and time-varying models—highlights the complexity of assessing SRI fund performance. Future research may benefit from integrating more granular ESG metrics and exploring the role of active versus passive management in sustainable investment strategies.

Despite the size of the market, the academic literature on the performance of assets still can't define a clear direction on the risk and return characteristics of those assets. For instance, Bauer et al. (2005) and Renneboog et al. (2011) find no significant relation between SRIs and performance, Borgers et al. (2015) and El Ghoul and Karoui (2017) declare SRIs as value destroying, while El Ghoul and Karoui (2022) and Nofsinger and Varma (2014) determine SRIs as value adding.

Consequently, the main topic of this research is to define whether a trading strategy in mutual funds based on past ESG ratings leads to consistent and significant excess returns compared to conventional investments.

The empirical analysis is rooted in the Carhart four-factor model, applied to a dataset of 4,000 U.S. mutual funds in two different grouping procedures, one over the period 2007-2023 and the other over the period 1990-2024. The funds are divided into equally weighted portfolios based on their top ten holdings' ESG scores' weighted average with a minimum threshold of 20% of known fund portfolio in the first analysis, while in the second also each fund sustainable investment techniques are considered and the minimum threshold of 20% of known fund portfolio is excluded. This grouping method allowed for a comparison of the returns between top-ranked funds and lower-ranked ones through the Carhart four factor model. Both analyses also distinguish between crisis and non-crisis periods, based on NBER defined recessions, ensuring that the impact of market downturns is captured. The methodology section of the thesis carefully outlines the statistical tools used, including regression analysis, the F-test for model significance, and the Ramsey RESET test for model robustness. These techniques help determine whether the performance of SRI funds is significantly driven by factors beyond the traditional market, size, value, and momentum variables.

The analysis yields several key findings. The top-performing funds, while exhibiting slight underperformance in crisis periods, do not generate statistically significant excess returns compared to their conventional counterparts. In contrast, lower-ranked funds show consistent negative alphas, particularly during crises, suggesting that they are more vulnerable to market downturns. These results suggest that mutual funds' returns are largely driven by systematic factors, with no significant evidence of alpha generated from ESG-based strategies. Interestingly, funds with sustainable investment strategies are not immune to the volatility associated with crises, and even the best-performing funds cannot fully avoid market declines during adverse economic conditions.

The thesis concludes by reaffirming that both SRI and conventional mutual fund portfolios do not necessarily lead to superior financial performance. The research underscores the importance of understanding the broader market forces at play while realizing that sustainability in investments is necessary for a livable future. Additionally, the findings indicate that the performance of mutual funds, whether sustainable or conventional, is primarily influenced by market dynamics, with little evidence of alpha creation that could be attributed to a fund's ESG strategy. As a result, the study adds to the continuing discussion concerning the risk-return characteristics of sustainable investments by highlighting the difficulties and complexities involved in incorporating ESG concerns into investment choices.

In the same way of Hamilton et al. (1993), Statman (2000), Goldreyer et al. (1999) and Schroder (2004), the results indicate that the market does not price social responsibility characteristics and that sustainable investments does not bring losses. It was rare to find a statistically significant difference between the α and even in that case, the value was small enough to be consumed by the investment expenses. Opposed to Goldreyer et al. (1999), by averaging α and computing the t-statistic for the difference, it is found that SRI mutual funds with investment screens underperform SRI mutual funds that don't apply them. Furthermore, in contrast with Borgers et al. (2015), Derwall et al. (2011) and Eccles et al. (2014), it is found that the α showed a negative performance for SRI mutual funds. Besides, in opposition with Galema et al. (2008), only in one case SRI resulted in lower HML factor, while usually it has a higher HML factor compared to conventional mutual funds. SRI in crisis periods did not outperform unlike Nofsinger and Varma (2014) but there's agreement in the fact that, overall, the alphas for SRI and conventional funds are insignificantly negative and not different from each other.

Looking ahead, future research could further explore the long-term performance of sustainable funds, especially as the global economy transitions towards a more sustainable model. This would involve a more granular analysis of specific ESG factors and their potential impact on various sectors and industries, as well as an investigation into the role of active versus passive fund management strategies in delivering superior ESG-aligned returns.

Chapter 2

Methodology

This section outlines the methodology used for conducting regression analysis in this study. The Ordinary Least Squares (OLS) regression framework is employed to examine the performance of mutual funds, utilizing the Carhart four-factor model. The statistical tests used to validate the regression results include the F-test, t-statistic, and Ramsey RESET test. These methods provide a robust approach to evaluating the significance and reliability of the regression estimates.

2.1 Carhart regression model

The Carhart regression model (Carhart 1997) is the product of a continuous development of portfolio theory. Starting from the Markowitz (1952) formulation of portfolio theory, the introduction of the assumption that investors can lend or borrow money at a risk-free rate led to the following CAPM equation by Sharpe (1964), Lintner (1965) and Mossin (1966):

$$R_{t,i} = RF_t + \beta_i(RM_t - RF_t) + \varepsilon_{i,t}$$

where $R_{t,i} - RF_t$ is defined as the return of asset i at time t in excess of the risk-free rate, β_i is the market beta of asset i and RM_t is the return on a market proxy. This model explains the return of an asset i as the sum of the return without risk and a risk-premium defined by the systematic risk beta times the market risk-premium.

Jensen (1968) improved the model by not making the regression line passing necessarily through the origin with a new intercept, called α, capturing the fact that, for a risky asset, the return can be higher or lower than the risk-premium associated to the market, thus associating a positive alpha with an asset that outperforms the market and a negative alpha with an asset that underperforms the market (Jensen 1968).

This finding paved the way for the introduction of additional independent variables in order to explain better the differences in excess returns.

In this empirical analysis, ten equally weighted mutual fund portfolios are created based on four thresholds of ESG-Scores, which are 5%, 10%, 25% and 50%, and sustainable investment screens. The first set of eight portfolios based on cutoff points of ESG scores are rebalanced every year to ensure continuity in the ESG rating of the portfolio and data availability for each month. The last two portfolios are not rebalanced during the period taken in consideration but returns are available for each month. The excess returns of such portfolios are assessed using the Fama and French (1993) three-factor model expanded with the Carhart (1997) momentum factor. This model controls for the impact of the market risk, the size factor, the book-to-market factor, and the momentum factor on returns:

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

where $R_{i,t}$ is the return on portfolio i, constructed as explained above, α_i is the monthly alpha for portfolio i, which indicates whether the portfolio has outperformed or underperformed the market after adjusting for risk factors, β_i measures the systematic risk of the portfolio, RM_t is the return in month t on a value-weighted market proxy, RF_t is the return in month t of a one-month treasury bill, s_i is the loading on the size factor, SMB_t is the difference in monthly return between a small and large-cap portfolio, h_i is the loading on the value factor, HML_t is the difference in return between a value and a growth portfolio, m_i is the loading on the momentum factor and MOM_t is the monthly return on a portfolio long on past one-year winners and short on past one-year losers. The momentum factor is designed to capture the risk due to the momentum found in stock returns by (Jeegadlesh and Titman 1993) while $\varepsilon_{t,i}$ refers to the idiosyncratic return component.

The above model is extended with the introduction of two dummy variables accounting for crisis and non-crisis periods:

$$R_{t,i} - RF_t = \alpha_{NC,i} D_{NC,t} + \alpha_{C,i} D_{C,t} + \beta_i(RM_t - RF_t) + s_i SMB_t + h_i HML_t + m_i MOM_t + \varepsilon_{t,i}$$

where $\alpha_{NC,i}$ is the non-crisis period monthly alpha for portfolio i, $D_{NC,t}$ is a dummy variable that takes the value of 1 if time t is defined as non-crisis period and 0 otherwise, $\alpha_{C,i}$ is the crisis period monthly alpha for portfolio i and $D_{C,t}$ is a dummy variable that takes the value of 1 if time t is defined as crisis period and 0 otherwise.

The data for RF_t, which is the 30-day Treasury Bill rate, $RM_t - RF_t$, SMB_t, HML_t and MOM_t are obtained from Kenneth French's (2012) web page. The data for crisis and non-crisis periods is given by the National Bureau of Economic Research (NBER) defined recession periods (National Bureau of Economic Research (NBER) 2023).

Moreover, the Carhart regression model (Carhart 1997) and its variant with dummy variables accounting for recession periods will also be used for mutual funds' individual regressions. The formula above keeps the same meaning except i won't refer to a portfolio anymore but a single mutual fund. The introduction of this approach stems from the fact that aggregate results of the previous regression might hide important features of funds such as the presence of outliers that consistently outperform its peers.

In the empirical analysis the Carhart regression model will be used with equally weighted portfolios of mutual funds and through individual regressions. In particular, the resulting α will measure the performance, with a positive alpha indicating that the portfolio has outperformed the market on a risk-adjusted basis and a negative alpha suggesting underperformance. When applying the Carhart model to an aggregate portfolio of funds, such as a set of ESG-focused or conventional mutual funds, the analysis helps in understanding the broad market trends and the systemic factors influencing these trends. This approach is beneficial for evaluating the overall performance of a market segment. Aggregating funds can smooth out individual fund anomalies, providing a clearer picture of the underlying factors that drive returns across a portfolio.

However, this method might obscure specific risks or opportunities that individual funds may present. For example, an aggregate analysis might not reveal a particular fund that consistently outperforms due to superior management or a unique strategy that aligns with market conditions not captured by the Carhart factors.

Conversely, applying the Carhart model to individual funds within the same ESG or conventional categories allows for a more granular analysis. This approach can highlight specific funds that either significantly outperform or underperform relative to their expected risk-adjusted returns based on market, size, value, and momentum factors.

Moreover, individual analysis can pinpoint how particular styles or strategies diverge from broader market behaviors, offering insights into risk management and the potential mispricing of assets. For instance, an individual fund that exhibits lower sensitivity to the momentum factor might be more stable during volatile market phases, a detail that aggregate analysis might miss.

After analyzing individual funds using the Carhart model, computing the mean performance across various funds can provide a comparative metric that balances the detail of individual analysis with broader, portfolio-level insights. This method can effectively highlight whether ESG funds, as a category, tend to perform better or worse than conventional funds after adjusting for market, size, value, and momentum effects. It also allows for the assessment of whether higher or lower ESG scores correlate with financial performance.

2.2 Statistical measures

In the context of Carhart regressions, various statistical tests are employed to ensure the robustness and appropriateness of the model specifications.

2.2.1 F-test

The F-test for "existence of regression" is employed in OLS regressions to test the joint significance of multiple explanatory variables. It specifically evaluates whether all regression coefficients, excluding the intercept, are simultaneously equal to zero. To test for the significance of the regression, we formulate the null hypothesis that all coefficients β, except for the intercept, are zero.

$$H_0 : \beta_2 = \beta_3 = \cdots = \beta_K = 0 \qquad H_1 : \text{at least one } \beta_i \text{ is different from 0 for } i = 2 \text{ to } K$$

Here, R and c are specified as:

$$R = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix} \quad \text{(a } (K-1) \times K \text{ matrix)}$$

$$c = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{(a } (K-1) \times 1 \text{ vector)}$$

The matrix R imposes $K-1$ constraints on K variables, excluding the intercept's coefficient. The F-test for the existence of the regression is then formulated as:

$$\text{F-test} = \frac{(R\hat{\beta} - c)'[R(X'X)^{-1}R']^{-1}(R\hat{\beta} - c)}{K-1} \div \frac{\hat{\varepsilon}'\hat{\varepsilon}}{N-K} \sim F(K-1, N-K)$$

The p-value is defined by $P(F \geq \hat{F})$. A larger F-statistic increases the likelihood of rejecting the null hypothesis, indicating that at least one of the coefficients significantly differs from zero. Conversely, a higher p-value suggests a greater likelihood of not rejecting the null hypothesis, indicating that the explanatory variables may not provide significant explanatory power over and above the intercept.

2.2.2 Univariate t-test

The t-statistic is commonly used in regression analysis to test whether an individual regression coefficient is statistically significant, evaluating the effect of an independent variable on the dependent variable.

The null and alternative hypotheses for a given coefficient β_j are defined as:

- H_0: $\beta_j = 0$, indicating no effect.
- H_1: $\beta_j \neq 0$, indicating a significant effect.

Consider a regression model with two coefficients:

$$y = \beta_1 + \beta_2 X + \varepsilon \quad H_0 : \beta_2 = 0 \quad H_1 : \beta_2 \neq 0$$

The t-statistic for β_2 is calculated as:

$$t = \frac{\hat{\beta}_2}{s.e.(\hat{\beta}_2)} \sim t(N-K)$$

where $\hat{\beta}_2$ is the estimated value of β_2, $s.e.(\hat{\beta}_2)$ is the standard error of $\hat{\beta}_2$, N is the number of observations, and K is the number of parameters (including the intercept).

Additionally, the t-statistic is utilized in the univariate analysis to assess significant differences between the means of two samples, which is particularly useful for data presumed to be normally distributed with unknown variances. The formula for the t-statistic in this context is:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

where $\bar{x}_1$ and $\bar{x}_2$ are the sample means of the first and second groups, respectively, s_1^2 and s_2^2 are the sample variances, and n_1 and n_2 are the sample sizes.

The p-value associated with the t-statistic helps determine whether to reject the null hypothesis. A small p-value implies that H_0 is likely false, indicating that the independent variable significantly affects the dependent variable in regression, or that there is a significant difference between the two sample means in a univariate test.

2.2.3 Ramsey RESET test

The Ramsey Regression Equation Specification Error Test (RESET) test introduces higher-order fitted values, such as squared and cubed terms of fitted values, into the regression model and examines their joint significance. The test statistic is computed as an F-test by comparing the original regression model with an augmented model that includes polynomial terms of the fitted values.

Define $w = [X, \hat{y}^2, \hat{y}^3]$, μ as the error term, $\psi = [\beta', \gamma_1, \gamma_2]'$, and $\hat{\psi} = (w'w)^{-1}w'y$. Also, $R = [\underbrace{0, \ldots, 0}_{2\times K}, I_2]$. The hypotheses are:

$$H_0 : \begin{cases} \gamma_1 = 0 \\ \gamma_2 = 0 \end{cases} \qquad H_1 : \text{unrestricted model, or at least one } \gamma \text{ different from } 0$$

The F-test is then given by:

$$\text{F-test} = \frac{\frac{[R\hat{\psi}-c]'[R(w'w)^{-1}R']^{-1}[R\hat{\psi}-c]}{2}}{\frac{\hat{\mu}'\hat{\mu}}{N-K-2}} \sim F(2, N-K-2)$$

A high p-value suggests that the null hypothesis cannot be rejected, indicating that the model is well-specified. A small p-value implies potential specification errors, such as omitted variables or incorrect functional form.

2.3 Potential Biases

The review of previous research highlighted a variety of potential biases and solutions. Concerning benchmark issues, the Carhart model is found to be helpful in solving them by controlling for biases such as small-cap and style biases, which caused ostensible outperformance or underperformance of SRI investments in previous studies. Indeed, by accounting for factors such as size, value, and momentum, it provides a more accurate assessment of mutual fund performance compared to conventional models (Bauer et al. 2005).

Regarding survivorship bias, which occurs when only surviving funds are analyzed, the findings suggests that it can skew performance results but this outcome can be mitigated by including dead funds in the analysis. Indeed, as pointed out by Brown et al. (1992), disregarding dead mutual funds leads to an overestimation of average performance.

Moreover, reverse causality, for which the direction of cause and effect is unclear between dependent and independent variables, can be addressed by introducing a longer lag between the dependent and independent variables (Eccles et al. 2014).

With respect to ESG scores, instead, ratings from different ESG-scores providers can yield differing results in the empirical analysis. Indeed, Berg et al. (2019) constructed a correlation matrix for ESG scores 2.1 for 924 firms in 2017 and found low correlation, with maximum value of 0.6 for Refinitiv with Moody's ESG and average between all rating agencies equal to 0.54 compared to the correlation value of around 0.99 across credit rating providers. The same matrix was computed for 2014 yielding similar results.

Further, greenwashing concerns, happening when companies falsely present themselves as sustainable, can be minimized by using data from long time periods, particularly data that predates the trend of sustainability.

Moreover, matching bias, which arises when pairing SRI and conventional funds for comparison, can be avoided by ensuring the matching occurs in different years to prevent year-specific effects from distorting the final sample. In some cases, the performance of SRI funds is influenced not by risk but by excess demand for these funds, as explained by the Fama-French factors. Separating the demand effects from risk factors is essential in such analyses.

Furthermore, fund turnover, which has been shown to negatively affect performance (Carhart 1997), should be closely examined when evaluating a fund's success. The skills

	SA MO	SA SP	SA RE	SA MS	MO SP	MO RE	MO MS	SP RE	SP MS	RE MS	Average
ESG	0.77	0.65	0.53	0.53	0.62	0.6	0.49	0.42	0.4	0.37	0.54
E	0.7	0.66	0.59	0.33	0.69	0.59	0.35	0.61	0.26	0.19	0.5
S	0.67	0.57	0.52	0.29	0.62	0.58	0.27	0.55	0.27	0.28	0.46
G	0.55	0.48	0.36	0.34	0.7	0.7	0.43	0.68	0.38	0.34	0.5

Figure 2.1: Correlations between ESG Ratings (2017)

Correlations between ESG ratings at the aggregate rating level (ESG) and at the level of the environmental dimension (E), the social dimension (S), and the governance dimension (G) for 924 firms. SA, SP, MO, RE and MS are short for Sustainalytics, S&P Global, Moody's ESG, Refinitiv and MSCI, respectively.

of mutual fund managers also play a crucial role in determining fund performance, making this an important factor to consider. SRI funds may exhibit industry bias, either favoring or avoiding certain sectors, which can distort performance comparisons with conventional funds. Finally, analyzing quarterly fund holdings can help control for differences in trading behavior, such as turnover and portfolio characteristics like firm capitalization, between SRI and conventional funds, which may influence overall performance (Nofsinger and Varma 2014).

This study addresses benchmark issues through the employment of the Carhart regression model to assess funds' excess returns, survivorship bias through the inclusion of dead mutual funds in the analysis, and reverse causality concerns by doing the matching procedure as early as 1990.

Chapter 3

Data

The sample of financial and ESG rating data for the US was retrieved from the LSEG Refinitiv and Datastream databases. The dataset consists of a random sample of 4000 domestic US mutual funds, both active and dead in order to avoid survivorship bias from the LSEG Refinitiv database, obtained by filtering available mutual funds for country of origin and country of exchange being the United States. Moreover, to find Lipper funds, for which time series data is available, the database was restricted to mutual fund classified with Lipper Classifications as either Equity US and Equity US Income. The dataset obtained was composed by 72 dead and 3928 active mutual funds. In order to group funds into sustainability categories, since the Refinitiv database registers time series of ESG scores only for stocks, the time series of the top 10 holdings for each mutual fund was downloaded, then an ESG score was attributed to each holding. ESG ratings were available in the databases for the stocks of 3944 funds, while an ulterior requirement of having at least 20% of the allocation for each fund further reduces the dataset size to 3736 funds. To each of these funds a score was assigned, computing the weighted average of the holdings' ESG scores with respect to their allocation in the fund portfolio and year. The time-series data offered by the databases span from January 1995 to December 2024 but the stated limitations restrict the timeframe of the current study to seventeen years, from 2007 to 2023.

Moreover, Refinitiv provides static data to represent sustainable investment techniques applied by mutual funds. After gathering the data, a sample of 211 mutual funds that applied at least one sustainable srategy was obtained and the same procedure to assign an ESG Combined score as above was applied, resulting in 149 sustainable mutual funds with ESG investment techniques 3.1. The matching procedure is a combination of that of (Bauer et al. 2005), based on fund age and fund size, and (Bollen 2007), based on fund launch date. For what concerns the launch date, the matched fund must be created within two years of the launch of the corresponding SRI fund, ensuring that the funds experience similar life-cycle effects and macroeconomic time-series effects (Bollen 2007). Regarding the total net asset value, the matched fund's total net assets differ from its corresponding SRI fund of at most 10%. Concerning fund age, the match is exact. Given those constraints, a matching set of 880 conventional mutual funds was built from the remaining data.

Table 3.1: Sustainable Investment Strategies and Number of Funds

Sustainable Investment Strategies	**Number of Funds**
Responsible Investments	211
ESG-Environmental	120
Positive Screening-Thematic	25
Positive Screening-Best in Class	24
Positive Screening-Positive Tilt	2
Negative Screening-ex Alcohol or Drugs	52
Negative Screening-ex Tobacco	71
Negative Screening-ex Weapons	74
Negative Screening-ex Adult Entertainment	63
Negative Screening-ex Fossil Energy	44
Negative Screening-ex GMO	18
Negative Screening-ex Nuclear	32
Negative Screening-ex Other	43
ESG-Governance	115
ESG-Social	128
Impact Investing-SDGs	15

Concerning the investment screens, Refinitiv defines them as follows: Responsible Investment identifies funds that include ESG, SRI, Positive/Negative Screening, Impact investing and or Religious criteria in their overall screening process; ESG-Environmental identifies funds that include environmental criteria in their overall screening process; Positive Screening-Thematic is an approach which invests in sustainable themes such as clean water, climate change, low carbon, low pollution, innovation and more; Positive Screening-Best in Class is an approach which identifies leading sustainable companies in a certain peer group which is not necessarily noted as sustainable; Positive Screening-

Positive Tilt is a technique that overweights leading companies compared to a benchmark; Negative Screening-ex Alcohol or Drugs identifies funds that exclude companies who are involved in the production or distribution of alcohol or drugs like cannabis from their investment universe; Negative Screening-ex Tobacco identifies funds that exclude companies who are involved in the production or distribution of tobacco from their investment universe; Negative Screening-ex Weapons identifies funds that exclude companies who are involved in the production of civilian and military weapons and firearms from their investment universe; Negative Screening-ex Fossil Energy identifies funds that exclude companies who are involved in the production or distribution of fossil energy from their investment universe. Fossil energy includes brown coal, stone coal, natural gas, mineral oil, thermal coal, oil sands and more. This may also include the producers of drilling equipment or equipment for refineries and plants; Negative Screening-ex GMO identifies funds that exclude companies who are involved in the production, distribution or use of genetically modified organisms from their investment universe; Negative Screening-ex Nuclear identifies funds that exclude companies who are involved in the production of nuclear power, nuclear power plants or uranium mining from their investment universe. This may also include the producers of parts for nuclear plants or other activities related to nuclear power; Negative Screening-ex Other identifies funds that excludes companies who are involved in the production or distribution of a segment which is currently not available as an exclusion segment; ESG-Governance identifies funds that include Governance criteria in their overall screening process; ESG-Social identifies funds that include social criteria in their overall screening process; Impact Investing- Sustainable Development Goals identifies funds that invest in companies that strive to have a positive contribution to the achievement of the UN sustainable development goals as part of the agenda 2030.

3.1 Refinitiv ESG Combined Score

The Refinitiv ESG Combined Score (ESGC) is a comprehensive measure designed to evaluate a company's Environmental, Social, and Governance (ESG) performance while integrating the impact of ESG-related controversies (Refinitiv LSEG 2024). The methodology ensures comparability, transparency, and industry relevance, allowing investors to

Table 3.2: Sample distribution, funds' total net assets and ESG combined score statistics based on ESG ranking

Fund Category	**# Funds**	**ESGC Score**	**TNA**		
Full Sample (2007 - 2023)			2007	2015	2023
Top 5% ESG Funds	807	62.139	$6.63 \cdot 10^{10}$	$2.395 \cdot 10^{12}$	$3.35 \cdot 10^{12}$
Top 10% ESG	1340	57.696	$1.01 \cdot 10^{11}$	$4.394 \cdot 10^{12}$	$6.40 \cdot 10^{12}$
Top 25% ESG	2272	51.706	$2.77 \cdot 10^{11}$	$9.453 \cdot 10^{12}$	$1.35 \cdot 10^{13}$
Top 50% ESG	3006	48.667	$4.37 \cdot 10^{11}$	$1.981 \cdot 10^{13}$	$3.00 \cdot 10^{13}$
Bottom 50% ESG	3131	48.667	$6.97 \cdot 10^{11}$	$1.361 \cdot 10^{13}$	$2.46 \cdot 10^{13}$
Bottom 25% ESG	1998	48.667	$2.02 \cdot 10^{11}$	$2.395 \cdot 10^{13}$	$4.11 \cdot 10^{13}$
Bottom 10% ESG	879	47.464	$4.06 \cdot 10^{11}$	$2.907 \cdot 10^{13}$	$4.83 \cdot 10^{13}$
Bottom 5% ESG	481	44.384	$4.13 \cdot 10^{11}$	$3.104 \cdot 10^{13}$	$5.12 \cdot 10^{13}$
Sustainable vs Conventional (1990 - 2024)					
Sustainable Funds	149				
Conventional Funds	880				

assess a company's sustainability practices objectively (Refinitiv LSEG 2024).

The ESG score is derived from company-reported data across three main pillars: Environmental, Social, and Governance (Refinitiv LSEG 2024). The Environmental pillar includes factors such as emissions reduction, energy efficiency, water use, and environmental product innovation (Refinitiv LSEG 2024). The Social pillar incorporates measures such as human rights policies, workforce diversity, labor practices, and product responsibility (Refinitiv LSEG 2024). The Governance pillar assesses corporate governance structures, shareholder rights, business ethics, and transparency (Refinitiv LSEG 2024). The ESG score is calculated using a weighted system where each company is benchmarked against its industry peers for environmental and social factors or its country of incorporation for governance factors (Refinitiv LSEG 2024).

To address industry-specific relevance, Refinitiv applies an ESG Magnitude Matrix that assigns weights to each ESG factor based on its materiality to a given industry (Refinitiv LSEG 2024). This ensures that industries with higher environmental impact, such as energy and mining, receive greater weight on environmental metrics, while governance metrics are benchmarked at the country level (Refinitiv LSEG 2024). The inclusion of materiality ensures that ESG factors are evaluated in a manner that reflects their relative significance within each industry (Refinitiv LSEG 2024).

The ESG Controversy Score is calculated based on media reports and public records on twenty-three types of ESG-related controversies, including environmental violations such as pollution and illegal waste disposal, social issues such as human rights violations and labor disputes, and governance concerns such as corruption and executive misconduct (Refinitiv LSEG 2024). These controversies are assigned severity weights based on the company's market capitalization, as larger companies naturally attract more media attention (Refinitiv LSEG 2024). The severity weighting mechanism ensures that controversy scores are adjusted appropriately to account for discrepancies in media exposure across different firm sizes (Refinitiv LSEG 2024).

The final ESGC score integrates both the ESG score and the ESG Controversy Score (Refinitiv LSEG 2024). If a company has no significant controversies, the ESGC score remains equal to the ESG score (Refinitiv LSEG 2024). However, if a company is involved in controversies, the ESGC score is adjusted downward based on the severity of the issues (Refinitiv LSEG 2024). The ESGC score is calculated as the average of the ESG score and the ESG Controversy Score (Refinitiv LSEG 2024). If the controversy score is lower than the ESG score, the ESGC score is adjusted to reflect the controversy impact (Refinitiv LSEG 2024). This adjustment ensures that companies with persistent or severe controversies do not receive inflated ESG scores (Refinitiv LSEG 2024).

The final ESG and ESGC scores are available in percentile rankings ranging from zero to one hundred and letter grades from D- to A+ (Refinitiv LSEG 2024). This simplifies interpretation, allowing investors to quickly assess a company's sustainability performance relative to its peers (Refinitiv LSEG 2024). The percentile ranking system provides a straightforward means of comparison across companies and industries, while the letter grading system enhances accessibility for a broader range of users (Refinitiv LSEG 2024).

The Refinitiv ESG Combined Score (ESGC) provides a holistic evaluation of a company's sustainability practices, incorporating both reported ESG performance and real-world controversies (Refinitiv LSEG 2024). By adjusting for industry materiality and controversy impacts, the ESGC score enhances comparability and reliability, making it a valuable tool for investors seeking to integrate ESG considerations into their decision-making process (Refinitiv LSEG 2024). The methodology ensures that ESG scores are not only based on self-reported data but also reflect a company's actual impact as perceived through external assessments. This dual approach strengthens the robustness of the ESGC score and reinforces its role as a key metric in sustainable investment strategies (Refinitiv LSEG 2024).

Chapter 4

Empirical Analysis

This section contains the results of the regressions applied to mutual funds ranked from the most to the least sustainable by their weighted average ESG combined score, obtained from the Refinitiv LSEG database, with respect to the percentage of the known portfolio ESG scores. The lower threshold is 20%, meaning that for every mutual fund in the analysis, at least 20% of the fund portfolio is known, in terms of both holdings and the holdings' ESG data. Given this limitation, the only available ESG data concern the period from January 2007 to December 2023. Each year from 2007 to 2023, an equally weighted portfolio of mutual funds is created for every ESG score threshold. Then there's the comparison of regressions of top and bottom ESG ranked mutual funds at 5%, 10%, 25%, 50% levels, considering both a full period regression and a regression with two dummies, accounting for both crisis and non-crisis periods as robustness check. The final number of mutual funds for each level is 807 for the top 5%, 1340 for the top 10%, 2272 for the top 25%, 3006 for the top 50%, 3131 for the bottom 50%, 1998 for the bottom 25%, 879 for the bottom 10% and 481 for the bottom 5%. Table 3.2 gives an idea of the size of mutual funds. Moreover, since the LSEG Refinitiv database registers funds who apply sustainability related strategies, such as positive and negative screening, impact finance and ESG related investments, by gathering all the funds that apply at least one of these screens, a sample of 149 sustainable and 880 conventional mutual funds was created. To the latter the same empirical analysis is applied, but this time starting from January of 1990 and ending in December of 2024. The analysis was conducted using Python and Matlab programming languages.

4.1 Comparative analysis based on ESG ratings

4.1.1 Performance of the top and bottom 50% ESG-rated Mutual Funds

Top 50% Mutual Funds

	Full period			**Crisis and Non-Crisis Periods**	
Parameter	Coefficient (β)	Std. Error	Parameter	Coefficient (β)	Std. Error
α	0.0113	0.0495	Crisis Period	0.1218	0.1349
	[0.2589]			[0.9029]	
			Non-Crisis Period	-0.0043	0.0471
				[-0.0907]	
$M - Rf$	0.9430***	0.0099	$M - Rf$	0.9454***	0.0103
	[95.0308]			[91.6548]	
SMB	-0.0026	0.0182	SMB	-0.0046	0.0184
	[-0.1401]			[-0.2518]	
HML	0.1000***	0.0137	HML	0.1018***	0.0138
	[7.3223]			[7.3636]	
MOM	-0.0263**	0.0122	MOM	-0.0242*	0.0125
	[-2.1522]			[-1.9349]	

Table 4.1: Performance of an equally weighted portfolio composed of the top 50% ESG-rated mutual funds with Carhart regression model. Full period regression tests: $R^2 = 0.9829$; F Test p-value = 0.0000; Ramsey RESET Test p-value = 0.8836. Crisis and non crisis period regression tests: $R^2 = 0.9829$; F Test p-value = 0.0000; Ramsey RESET Test p-value = 0.7446.

Notes: t-statistics are reported in brackets. Significance levels are denoted as follows: *** p<0.01, ** p<0.05, * p<0.10.

The Carhart four-factor regression on the top 50% mutual funds equally weighted portfolio in table 4.1 shows that market exposure is the key driver of returns, with a beta close to one and strong statistical significance. The portfolio has a considerable and consistent tilt toward value equities, as seen by the significantly positive HML coefficient, whilst the SMB component is statistically insignificant, showing no preference for small- or large-cap firms. The momentum component is negative and significant, indicating an anti-momentum impact in which the portfolio underperforms equities with high historical returns. The addition of crisis and non-crisis dummies yields no significant coefficients, showing that portfolio performance is not significantly different across market circumstances. The model explains nearly all return variation with $R^2 = 98.29\%$, and the Ramsey RESET test confirms a well-specified regression. These findings suggest that the portfolio's returns are predominantly systematic, with no evidence of crisis-driven abnormal performance.

Bottom 50% Mutual Funds

Full period			**Crisis and Non-Crisis Periods**		
Parameter	Coefficient (β)	Std. Error	Parameter	Coefficient (β)	Std. Error
α	-0.0089 [-0.1782]	0.0500	Crisis Period	0.1397 [0.9023]	0.1549
			Non-Crisis Period	-0.0298 [-0.5512]	0.0541
$M - Rf$	0.9893*** [86.8141]	0.0114	$M - Rf$	0.9925*** [83.8389]	0.0118
SMB	0.2325*** [11.1138]	0.0209	SMB	0.2297*** [10.8870]	0.0211
HML	-0.0293* [-1.8692]	0.0157	HML	-0.0268* [-1.6900]	0.0159
MOM	-0.0298** [-2.1228]	0.0141	MOM	-0.0269* [-1.8779]	0.0143

Table 4.2: Performance of an equally weighted portfolio composed of the bottom 50% ESG-rated mutual funds with Carhart regression model. Full period regression tests: $R^2 = 0.9806$; F Test p-value = 0.0000; Ramsey RESET Test p-value = 0.9244. Crisis and non crisis period regression tests: $R^2 = 0.9807$; F Test p-value = 0.0000; Ramsey RESET Test p-value = 0.9016.

Notes: t-statistics are reported in brackets. Significance levels are denoted as follows: *** p<0.01, ** p<0.05, * p<0.10.

The results of the Carhart four-factor regression on an equally weighted portfolio of mutual funds in table 4.2 reveal that the portfolio's returns are predominantly driven by market exposure, with a beta close to one and high statistical significance. While the HML factor is negative but only marginally significant, showing a slight predisposition for growth over value companies, the SMB factor is constantly positive and significant, indicating a strong tilt toward small-cap securities. A negative and statistically significant momentum factor (MOM) suggests an anti-momentum impact, in which the portfolio often performs worse than previous winners. The performance of the portfolio does not significantly change between crisis and non-crisis periods, as seen by the lack of statistically significant coefficients produced by including crisis and non-crisis dummies. The robustness of the regression is confirmed by the strong R2 values, which show that the model accounts for almost all of the variation in returns, and the Ramsey RESET test, which indicates no specification errors. Overall, the findings suggest that the portfolio's returns are largely systematic and do not exhibit abnormal performance across market cycles.

Top 50% - Bottom 50% Mutual Funds

Full period			**Crisis and Non-Crisis Periods**		
Parameter	Coefficient (β)	Std. Error	Parameter	Coefficient (β)	Std. Error
α	-0.0690*	0.0400	Non-Crisis Period	-0.0631	0.0434
	[-1.7259]			[-1.4559]	
			Crisis Period	-0.1111	0.1242
				[-0.8944]	
$M - Rf$	-0.0453***	0.0091	$M - Rf$	-0.0462***	0.0095
	[-4.9652]			[-4.8652]	
SMB	-0.2313***	0.0167	SMB	-0.2305***	0.0169
	[-13.8147]			[-13.6197]	
HML	0.1324***	0.0125	HML	0.1317***	0.0127
	[10.5560]			[10.3508]	
MOM	0.0038	0.0112	MOM	0.0030	0.0115
	[0.3354]			[0.2566]	

Table 4.3: Performance of an equally weighted portfolio long on the top 50% mutual funds and short on the bottom 50% mutual funds with Carhart regression model. Full period regression tests: $R^2 = 0.6512$; F Test p-value = 0.0000; Ramsey RESET Test p-value = 0.6404. Crisis and non crisis period regression tests: $R^2 = 0.6514$; F Test p-value = 0.0000; Ramsey RESET Test p-value = 0.6641.

Notes: t-statistics are reported in brackets. Significance levels are denoted as follows: *** p<0.01, ** p<0.05, * p<0.10.

Both the top and bottom 50% mutual fund portfolios' α do not exhibit any notable abnormal returns beyond factor exposures, according to a comparison of their returns. As a result, it is impossible to identify which of the two mutual fund sets has larger anomalous returns at the 50% threshold. The outcome is different, though, when looking at an investment in a portfolio that goes long on top-ranked funds and short on low-ranked ones in table 4.3. Specifically, this portfolio produces a slightly significant negative α, indicating that the mutual funds with the lowest 50% ESG rating may perform better than their counterparts. When recession times are taken into account, this does not occur, indicating that there are no exceptional returns. The HML factor became highly statistically significant in this portfolio, proving that both SRI and conventional mutual funds prefer to invest in value over growth stock, opposed to the results yielded previously by the two individual portfolios. Moreover, from table $B.7$ in the appendix, the t-statistic on the difference of mean α is not statistically significant, proving once again that returns of both sustainable and conventional mutual funds are similar.

4.1.2 Performance of the top and bottom 25% ESG-rated Mutual Funds

Top 25% Mutual Funds

Full period			**Crisis and Non-Crisis Periods**		
Parameter	Coefficient (β)	Std. Error	Parameter	Coefficient (β)	Std. Error
α	-0.0581	0.0465	Crisis Period	-0.2147	0.1440
	[-1.2484]			[-1.4913]	
			Non-Crisis Period	-0.0361	0.0503
				[-0.7172]	
$M - Rf$	0.9374***	0.0106	$M - Rf$	0.9340***	0.0110
	[88.4044]			[84.8363]	
SMB	-0.0178	0.0195	SMB	-0.0148	0.0196
	[-0.9129]			[-0.7557]	
HML	0.1206***	0.0146	HML	0.1180***	0.0148
	[8.2682]			[7.9976]	
MOM	-0.0275**	0.0131	MOM	-0.0306**	0.0133
	[-2.1054]			[-2.2940]	

Table 4.4: Performance of an equally weighted portfolio composed of the top 25% ESG-rated mutual funds with Carhart regression model. Full period regression tests: $R^2 = 0.9803$; F Test p-value = 0.0000; Ramsey RESET Test p-value = 0.0612. Crisis and non crisis period regression tests: $R^2 = 0.9805$; F Test p-value = 0.0000; Ramsey RESET Test p-value = 0.0847.

Notes: t-statistics are reported in brackets. Significance levels are denoted as follows: *** p<0.01, ** p<0.05, * p<0.10.

With a beta around one and high statistical significance, the Carhart four-factor regression on an equally weighted mutual fund portfolio in table 4.4 indicates an important dependence on market risk. The statistically insignificant SMB coefficient shows that there is not an apparent tilt in the portfolio toward small- or large-cap firms. However, as indicated by a significantly positive HML coefficient, the portfolio exhibits a strong and consistent predisposition for value equities. Furthermore, the momentum component is statistically significant and negative, suggesting an anti-momentum effect in which the portfolio often performs worse than previous winners. The lack of significant coefficients from the addition of crisis and non-crisis dummies suggests that portfolio performance is consistent across a range of market circumstances. Almost all return variance is explained by the model, however the Ramsey RESET test suggests a mild specification issue, which may warrant further investigation into potential nonlinearities or omitted variables. Overall, the findings indicate that the portfolio's returns are primarily driven by systematic risk factors rather than time-dependent effects.

Bottom 25% Mutual Funds

Full period			**Crisis and Non-Crisis Periods**		
Parameter	Coefficient (β)	Std. Error	Parameter	Coefficient (β)	Std. Error
α	-0.1198*	0.0659	Crisis Period	-0.4708**	0.2030
	[-1.8170]			[-2.3190]	
			Non-Crisis Period	-0.0704	0.0709
				[-0.9936]	
$M - Rf$	1.0406***	0.0150	$M - Rf$	1.0329***	0.0155
	[69.2567]			[66.5464]	
SMB	0.4038***	0.0276	SMB	0.4104***	0.0277
	[14.6356]			[14.8339]	
HML	-0.0735***	0.0207	HML	-0.0794***	0.0208
	[-3.5564]			[-3.8167]	
MOM	0.0095	0.0185	MOM	0.0026	0.0188
	[0.5111]			[0.1393]	

Table 4.5: Performance of an equally weighted portfolio composed of the bottom 25% ESG-rated mutual funds with Carhart regression model. Full period regression tests: $R^2 = 0.9712$; F Test p-value = 0.0000; Ramsey RESET Test p-value = 0.2650. Crisis and non crisis period regression tests: $R^2 = 0.9717$; F Test p-value = 0.0000; Ramsey RESET Test p-value = 0.4642.

Notes: t-statistics are reported in brackets. Significance levels are denoted as follows: *** p<0.01, ** p<0.05, * p<0.10.

The regression findings in table 4.5 demonstrate that market exposure plays a dominating role, with a beta slightly over one, suggesting the portfolio is more volatile than the market. A strong and statistically significant SMB coefficient indicates a preference for small-cap equities, whereas a negative and significant HML coefficient indicates that the portfolio favours growth stocks to value stocks. Unlike in other portfolios, momentum does not appear to have a substantial impact on results. The crisis and non-crisis period analysis shows that while the portfolio does not experience statistically significant shifts in performance during normal market conditions, it does exhibit significant underperformance in crisis periods, as indicated by the negative and highly significant crisis dummy (α). Despite this, the regression model explains nearly all return variation, and the Ramsey RESET test confirms a well-specified model.

Top 25% - Bottom 25% Mutual Funds

Full period			**Crisis and Non-Crisis Periods**		
Parameter	Coefficient (β)	Std. Error	Parameter	Coefficient (β)	Std. Error
α	-0.0275 [-0.4021]	0.0684	Non-Crisis Period	-0.0543 [-0.7331]	0.0740
			Crisis Period	0.1629 [0.7681]	0.2121
$M - Rf$	-0.1022*** [-6.5495]	0.0156	$M - Rf$	-0.0980*** [-6.0439]	0.0162
SMB	-0.4177*** [-14.5849]	0.0286	SMB	-0.4213*** [-14.5803]	0.0289
HML	0.1973*** [9.1928]	0.0215	HML	0.2005*** [9.2266]	0.0217
MOM	-0.0367* [-1.9091]	0.0192	MOM	-0.0330* [-1.6810]	0.0196

Table 4.6: Performance of an equally weighted portfolio long on the top 25% mutual funds and short on the bottom 25% mutual funds with Carhart regression model. Full period regression tests: $R^2 = 0.6717$; F Test p-value = 0.0000; Ramsey RESET Test p-value = 0.2565. Crisis and non crisis period regression tests: $R^2 = 0.6732$; F Test p-value = 0.0000; Ramsey RESET Test p-value = 0.2198.

Notes: t-statistics are reported in brackets. Significance levels are denoted as follows: *** p<0.01, ** p<0.05, * p<0.10.

A comparative analysis of the Carhart four-factor regression for the bottom 25% and top 25% ESG-ranked equally weighted mutual funds portfolios highlights key differences in factor exposures and performance characteristics. Performance across market conditions remains stable for both portfolios during crisis periods, but the bottom 25% funds significantly underperform in non-crisis periods, whereas the top 25% funds do not exhibit such a pattern. Both models exhibit strong explanatory power, though the top 25% funds model presents mild specification concerns. Overall, the findings suggest that the top 25% funds generate returns more efficiently through value investing strategies, while the bottom 25% funds rely more on small-cap and growth stocks and underperform during crisis periods. Additionally, considering a strategy in which the portfolio is formed by going long on the top 25% and short on the bottom 25%, the Carhart regression does not sustain the claim of a marginally statistically significant negative α. Indeed, there's a negative α for both full and recession periods regressions, but it is not significant, while the other factors stand out in the unique portfolio. Once again, the performances of sustainable and conventional funds in the individual regressions' means in table B.5 do not differ much given a t-statistic not statistically significant.

4.1.3 Performance of the top and bottom 10% ESG-rated Mutual Funds

Top 10% Mutual Funds

Full period			**Crisis and Non-Crisis Periods**		
Parameter	Coefficient (β)	Std. Error	Parameter	Coefficient (β)	Std. Error
α	-0.1071	0.0686	Crisis Period	-0.4158*	0.2117
	[-1.5623]			[-1.9640]	
			Non-Crisis Period	-0.0637	0.0739
				[-0.8624]	
$M - Rf$	0.9465***	0.0156	$M - Rf$	0.9397***	0.0162
	[60.5468]			[58.0521]	
SMB	-0.0144	0.0287	SMB	-0.0086	0.0288
	[-0.5012]			[-0.2975]	
HML	0.1052***	0.0215	HML	0.1000***	0.0217
	[4.8919]			[4.6115]	
MOM	-0.0089	0.0193	MOM	-0.0149	0.0196
	[-0.4610]			[-0.7605]	

Table 4.7: Performance of an equally weighted portfolio composed of the top 10% ESG-rated mutual funds with Carhart regression model. Full period regression tests: R^2 = 0.9583; F Test p-value = 0.0000; Ramsey RESET Test p-value = 0.0302. Crisis and non crisis period regression tests: R^2 = 0.9588; F Test p-value = 0.0000; Ramsey RESET Test p-value = 0.0480.

Notes: t-statistics are reported in brackets. Significance levels are denoted as follows: *** p<0.01, ** p<0.05, * p<0.10.

Using an equally weighted portfolio of the top 10% mutual funds in table 4.7, the Carhart four-factor regression provides important information on factor exposures and performance attributes. Market risk is the primary driver of returns, according to the regression results, which have a beta near one with high statistical significance. As the statistically insignificant SMB coefficient shows, the portfolio does not show a substantial tilt toward small- or large-cap stocks. Nonetheless, with a significantly positive HML coefficient, a clear preference for value stocks is evident. The momentum element does not significantly contribute to the explanation of returns, in contrast to other portfolios examined. The portfolio has a negative significant α during crisis periods, whereas in normal market conditions, according to the examination of crisis and non-crisis periods, it has no abnormal returns. The Ramsey RESET test indicates the presence of potential nonlinear effects or omitted variables, suggesting a possible specification error even though the model explains a significant amount of return variation. These results imply that market exposure and a preference for value stocks are the main ways in which the top 10 mutual funds produce returns, with little to no dependence on size or momentum.

Bottom 10% Mutual Funds

Full period			**Crisis and Non-Crisis Periods**		
Parameter	Coefficient (β)	Std. Error	Parameter	Coefficient (β)	Std. Error
α	-0.1676*	0.0906	Crisis Period	-0.6837**	0.2788
	[-1.8501]			[-2.4528]	
			Non-Crisis Period	-0.0951	0.0973
				[-0.9773]	
$M - Rf$	1.0345***	0.0207	$M - Rf$	1.0232***	0.0213
	[50.0845]			[48.0108]	
SMB	0.5365***	0.0379	SMB	0.5462***	0.0380
	[14.1474]			[14.3809]	
HML	-0.1159***	0.0284	HML	-0.1245***	0.0286
	[-4.0787]			[-4.3606]	
MOM	0.0486*	0.0255	MOM	0.0385	0.0258
	[1.9067]			[1.4914]	

Table 4.8: Performance of an equally weighted portfolio composed of the bottom 10% ESG-rated mutual funds with Carhart regression model. Full period regression tests: R^2 = 0.9485; F Test p-value = 0.0000; Ramsey RESET Test p-value = 0.2420. Crisis and non crisis period regression tests: R^2 = 0.9495; F Test p-value = 0.0000; Ramsey RESET Test p-value = 0.3796.

Notes: t-statistics are reported in brackets. Significance levels are denoted as follows: *** p<0.01, ** p<0.05, * p<0.10.

The Carhart four-factor regression on an equally weighted portfolio of the least sustainable 10% mutual funds in table 4.8 demonstrates important exposures that distinguish these funds from better-performing portfolios. The regression findings show that market exposure is the primary driver of returns, with a beta slightly higher than one, implying that the portfolio is more volatile than the market. The portfolio has a strong and a significant tilt toward small-cap equities, as shown by the particularly significant SMB coefficient. However, it also exhibits a strong preference for growth equities, as seen by the significantly negative HML coefficient. The momentum component seems to be positive but only slightly significant, indicating that the portfolio may benefit from prior winners. An important discovery from the crisis and non-crisis periods analysis is that while the portfolio does not experience statistically significant shifts in performance in normal market conditions, it does exhibit significant underperformance during crisis periods, as indicated by the negative and highly significant crisis dummy. Despite these weaknesses, the model explains nearly all return variation, and the Ramsey RESET test confirms a well-specified model. These data imply that the lowest 10% funds, while considerably exposed to small-cap growth equities, tend to suffer in stable market conditions. This might be attributed to increased volatility and worse performance consistency.

Top 10% - Bottom 10% Mutual Funds

Full period			**Crisis and Non-Crisis Periods**		
Parameter	Coefficient (β)	Std. Error	Parameter	Coefficient (β)	Std. Error
α	-0.0287 [-0.2913]	0.0985	Non-Crisis Period	-0.0573 [-0.5370]	0.1067
			Crisis Period	0.1747 [0.5717]	0.3056
$M - Rf$	-0.0870*** [-3.8753]	0.0225	$M - Rf$	-0.0826*** [-3.5338]	0.0234
SMB	-0.5471*** [-13.2695]	0.0412	SMB	-0.5509*** [-13.2310]	0.0416
HML	0.2242*** [7.2589]	0.0309	HML	0.2276*** [7.2712]	0.0313
MOM	-0.0572** [-2.0647]	0.0277	MOM	-0.0532* [-1.8804]	0.0283

Table 4.9: Performance of an equally weighted portfolio long on the top 10% mutual funds and short on the bottom 10% mutual funds with Carhart regression model. Full period regression tests: $R^2 = 0.5888$; F Test p-value = 0.0000; Ramsey RESET Test p-value = 0.8230. Crisis and non crisis period regression tests: $R^2 = 0.5898$; F Test p-value = 0.0000; Ramsey RESET Test p-value = 0.8113.

Notes: t-statistics are reported in brackets. Significance levels are denoted as follows: *** p<0.01, ** p<0.05, * p<0.10.

For what concerns the relative performance of SRI mutual funds with respect to non-SRI mutual funds, the regression exhibits a marginally significant negative alpha in crisis periods while non-SRI mutual funds are characterized by a significant negative alpha during both full period regression and crisis periods, suggesting that non-SRI mutual funds underperform SRI mutual funds at the 10% level.

However, by computing a regression of an equally weighted portfolio formed by going long on the top 10% funds and short on the bottom 10% funds, α is negative but not statistically significant, whereas the other factors maintain high statistical significance both on full and recession periods, suggesting that conventional and sustainable mutual funds apply very similar strategies. This is confirmed by the t-statistic of the difference in average alpha of individual regressions on both groups of funds in table B.3.

4.1.4 Performance of the top and bottom 5% ESG-rated Mutual Funds

Top 5% Mutual Funds

Full period			**Crisis and Non-Crisis Periods**		
Parameter	Coefficient (β)	Std. Error	Parameter	Coefficient (β)	Std. Error
α	-0.0962	0.0754	Crisis Period	-0.3548	0.2336
	[-1.2766]			[-1.5187]	
			Non-Crisis Period	-0.0593	0.0817
				[-0.7263]	
$M - Rf$	0.9435***	0.0169	$M - Rf$	0.9377***	0.0176
	[55.7110]			[53.1627]	
SMB	-0.0079	0.0310	SMB	-0.0032	0.0312
	[-0.2546]			[-0.1034]	
HML	0.1019***	0.0231	HML	0.0972***	0.0234
	[4.4157]			[4.1553]	
MOM	-0.0122	0.0209	MOM	-0.0179	0.0215
	[-0.5836]			[-0.8337]	

Table 4.10: Performance of an equally weighted portfolio composed of the top 5% ESG-rated mutual funds with Carhart regression model. Full period regression tests: $R^2 = 0.9544$; F Test p-value = 0.0000; Ramsey RESET Test p-value = 0.1308. Crisis and non crisis period regression tests: $R^2 = 0.9548$; F Test p-value = 0.0000; Ramsey RESET Test p-value = 0.1801.

Notes: t-statistics are reported in brackets. Significance levels are denoted as follows: *** p<0.01, ** p<0.05, * p<0.10.

The Carhart four-factor regression on an equally weighted portfolio of top 5% mutual funds provides valuable insights on factor exposures and performance characteristics. The regression results show that market risk is the dominant driver of returns, with a beta around one and a strong statistical significance. The portfolio does not have a substantial tilt toward small- or large-cap stocks, as demonstrated by the statistically insignificant SMB coefficient. However, there is a substantial preference for value equities, as reflected by a significantly positive HML coefficient. Unlike the other portfolios examined, the momentum component does not play a significant role in explaining results.

The examination of crisis and non-crisis time periods reveals no substantial changes in performance, implying that the portfolio acts similarly under different market situations. The model explains a significant percentage of return variance, and the Ramsey RESET test finds no specification errors, showing that the regression is well-specified. These data imply that the top 5% of mutual funds produce returns largely through market exposure and a preference for value equities, without significant dependence on size or momentum variables.

Bottom 5% Mutual Funds

	Full period			**Crisis and Non-Crisis Periods**	
Parameter	Coefficient (β)	Std. Error	Parameter	Coefficient (β)	Std. Error
α	-0.1030	0.1109	Crisis Period	-0.2489	0.3446
	[-0.9296]			[-0.7222]	
			Non-Crisis Period	-0.0822	0.1205
				[-0.6825]	
$M - Rf$	0.9748***	0.0249	$M - Rf$	0.9715***	0.0260
	[39.1444]			[37.3425]	
SMB	0.6859***	0.0456	SMB	0.6885***	0.0461
	[15.0417]			[14.9436]	
HML	-0.1467***	0.0339	HML	-0.1493***	0.0345
	[-4.3231]			[-4.3268]	
MOM	0.0403	0.0308	MOM	0.0371	0.0316
	[1.3093]			[1.1714]	

Table 4.11: Performance of an equally weighted portfolio composed of the bottom 5% ESG-rated mutual funds with Carhart regression model. Full period regression tests: $R^2 = 0.9301$; F Test p-value = 0.0000; Ramsey RESET Test p-value = 0.1412. Crisis and non crisis period regression tests: $R^2 = 0.9301$; F Test p-value = 0.0000; Ramsey RESET Test p-value = 0.1600.

Notes: t-statistics are reported in brackets. Significance levels are denoted as follows: *** p<0.01, ** p<0.05, * p<0.10.

The Carhart four-factor regression on an equally weighted portfolio of the bottom 5% of mutual funds identifies crucial performance characteristics. The findings show that market exposure is a significant driver of returns, with a beta close to one and strong statistical significance. A strong and extremely significant SMB coefficient shows an obvious preference for small-cap equities, whereas a negative and significant HML coefficient indicates an obvious lean towards growth stocks over value stocks. Unlike several higher-ranked portfolios, the momentum element has no substantial influence on returns. The examination of crisis and non-crisis periods reveals no substantial changes in performance, implying that the portfolio acts similarly under different market situations. The model explains a significant percentage of the return fluctuation, and the Ramsey RESET test shows no specification errors, suggesting a well-specified regression.

Top 5% - Bottom 5% Mutual Funds

Full period			**Crisis and Non-Crisis Periods**		
Parameter	Coefficient (β)	Std. Error	Parameter	Coefficient (β)	Std. Error
α	-0.0649	0.1344	Non-Crisis Period	-0.1161	0.1506
	[-0.4829]			[-0.7710]	
			Crisis Period	0.6586	0.4185
				[1.5737]	
$M - Rf$	-0.0300	0.0302	$M - Rf$	-0.0275	0.0307
	[-0.9954]			[-0.8972]	
SMB	-0.6913***	0.0553	SMB	-0.6969***	0.0553
	[-12.5074]			[-12.6093]	
HML	0.2503***	0.0411	HML	0.2549***	0.0410
	[6.0837]			[6.2100]	
MOM	-0.0512	0.0373	MOM	-0.0338	0.0389
	[-1.3737]			[-0.8678]	

Table 4.12: Performance of an equally weighted portfolio long on the top 5% mutual funds and short on the bottom 5% mutual funds with Carhart regression model. Full period regression tests: $R^2 = 0.5322$; F Test p-value = 0.0000; Ramsey RESET Test p-value = 0.7760. Crisis and non crisis period regression tests: $R^2 = 0.5381$; F Test p-value = 0.0000; Ramsey RESET Test p-value = 0.7252.

Notes: t-statistics are reported in brackets. Significance levels are denoted as follows: *** p<0.01, ** p<0.05, * p<0.10.

At the 5% level, conclusions can't be drawn on abnormal returns, since for both SRI and non-SRI funds, both for crisis and non crisis periods, the α are not significant.

Nevertheless, computing a regression on an equally weighted portfolio that goes long on sustainable and short on conventional mutual funds in table 4.12, the results suggest that both top and low-ranked funds maintain a negative beta on the SMB factor and a positive beta on the HML factor, as opposed to the regression on the low-ranked portfolio in table 4.11, suggesting that the combined portfolios are tilted towards big-cap and value stocks. This relation is maintained in the recession periods. Even in this case the t-statistic of the difference of the means of α in table B.1 is not statistically significant. However the average alpha, both in the full period and in recession periods, is negative.

4.2 Comparative analysis based on Sustainable Investment techniques

Mutual Funds with a sustainable investment strategy

Full period			**Crisis and Non-Crisis Periods**		
Parameter	Coefficient (β)	Std. Error	Parameter	Coefficient (β)	Std. Error
α	-0.0256	0.0400	Crisis Period	-0.0462	0.0425
	[-0.6400]			[-1.0872]	
			Non-Crisis Period	0.1544	0.1321
				[1.1690]	
$M - Rf$	0.9818***	0.0094	$M - Rf$	0.9842***	0.0095
	[104.9576]			[103.7541]	
SMB	0.0569***	0.0129	SMB	0.0548***	0.0130
	[4.3993]			[4.2197]	
HML	-0.0404***	0.0125	HML	-0.0382***	0.0126
	[-3.2294]			[-3.0288]	
MOM	-0.0172*	0.0091	MOM	-0.0154*	0.0092
	[-1.8817]			[-1.6679]	

Table 4.13: Performance of an equally weighted portfolio composed of the funds applying a sustainable investment strategy with Carhart regression model. Full period regression tests: $R^2 = 0.9692$; F Test p-value = 0.0000; Ramsey RESET Test p-value = 0.7470. Crisis and non crisis period regression tests: $R^2 = 0.9694$; F Test p-value = 0.0000; Ramsey RESET Test p-value = 0.4876.

Notes: t-statistics are reported in brackets. Significance levels are denoted as follows: *** p<0.01, ** p<0.05, * p<0.10.

An equally weighted mutual fund portfolio in table 4.13 was evaluated according to a Carhart four-factor regression analysis, which yielded important insights into the factors influencing portfolio returns for the whole sample period as well as during periods of crisis. With an exceptionally significant beta, the first regression, which covers the whole period, shows that the market excess return component is the main driver of mutual fund performance. While the HML factor is significant and negative, suggesting a bias for growth over value companies, the SMB factor is positive and statistically significant, suggesting a preference for small-cap equities. There is a slight negative impact from the momentum component. The alpha term, which denotes abnormal returns that cannot be explained by these factors, is noteworthy because it is insignificant, confirming that fund performance is mostly due to the systematic component rather than a result of active management skill. A well-specified model is indicated by the model's strong explanatory power and the Ramsey RESET test's inability to identify specification problems.

Crisis and non-crisis dummy variables are used in the second regression to evaluate how macroeconomic conditions affect fund performance. A negative but statistically insignificant alpha for the crisis dummy and no statistically significant alpha for the non-crisis dummy indicate that the portfolio does not show abnormal returns during either period. The market component continues to dominate, with a beta that is practically unchanged, confirming the strong correlation between mutual fund returns and the market as a whole. According to the full-period study, the SMB and HML variables continue to be significant and retain their signs. Although statistically weak, the momentum component is still negative, and the Ramsey RESET test once more shows no significant specification errors. Overall, the findings suggest that the portfolio's performance is primarily driven by systematic factors, with no evidence of significant abnormal returns during crisis or non-crisis periods. Given that factor exposures account for almost all return volatility, this emphasizes that challenge of consistently generating outperformance through active management.

Conventional Mutual Funds

Full period			**Crisis and Non-Crisis Periods**		
Parameter	Coefficient (β)	Std. Error	Parameter	Coefficient (β)	Std. Error
α	0.0195 [0.4739]	0.0411	Crisis Period	0.0108 [0.2466]	0.0437
			Non-Crisis Period	0.0957 [0.7034]	0.1360
$M - Rf$	1.0040*** [104.4596]	0.0096	$M - Rf$	1.0050*** [102.9013]	0.0098
SMB	0.3772*** [28.4046]	0.0133	SMB	0.3763*** [28.1423]	0.0134
HML	0.0589*** [4.5792]	0.0129	HML	0.0599*** [4.6130]	0.0130
MOM	0.0006 [0.0626]	0.0094	MOM	0.0014 [0.1433]	0.0095

Table 4.14: Performance of an equally weighted portfolio composed of the conventional funds matched with funds that have a sustainable investment strategy with Carhart regression model. Full period regression tests: $R^2 = 0.9729$; F Test p-value = 0.0000; Ramsey RESET Test p-value = 0.1732. Crisis and non crisis period regression tests: $R^2 = 0.9730$; F Test p-value = 0.0000; Ramsey RESET Test p-value = 0.1395.

Notes: t-statistics are reported in brackets. Significance levels are denoted as follows: *** p<0.01, ** p<0.05, * p<0.10.

The Carhart four-factor regression on an equally weighted portfolio of mutual funds in table 4.14 provides insights into the role of systematic factors in explaining portfolio performance. The first regression, which covers the whole sample period, shows that the market excess return component is the primary driver of returns, with a large beta. The SMB factor is positive and extremely significant, implying a strong preference for small-cap stocks. The HML factor is likewise positive and statistically significant, indicating a preference for value stocks. However, the momentum component does not appear to have a significant impact on returns; its coefficient is near to zero and statistically insignificant. The alpha element, which represents unexplained anomalous returns, is also insignificant, implying that factor exposures account for the majority of fund performance. The high R^2 value of 97.29% indicates that the model captures nearly all variations in fund returns, and the Ramsey RESET test suggests no specification errors.

The second regression uses dummy variables for crisis and non-crisis periods to compare performance across economic situations. The crisis dummy is positive but statistically insignificant, implying that there is no clear evidence of abnormal performance during financial downturns. The non-crisis dummy is equally insignificant, indicating that returns are stable across market conditions. The market beta stays close to one, indicating the portfolio's high reliance on market moves. The SMB and HML variables maintain their significance and signs, replicating the results of the full-period regression. The momentum factor remains insignificant, and the Ramsey RESET test does not reveal any misspecification issues. These data indicate that mutual fund returns are mostly influenced by systematic risk exposures, with no noteworthy abnormal performance during crisis and non-crisis periods. This supports the notion that active management does not generate consistent excess returns beyond what is captured by standard risk factors.

Importantly, α is statistically insignificant in both regressions, indicating that neither portfolio exhibits abnormal returns beyond what is explained by standard risk factors. The higher R^2 value for non-sustainable funds suggests a slightly better fit of the Carhart model in explaining these returns. Additionally, the Ramsey RESET test p-value is lower, implying that sustainable fund returns exhibit fewer signs of model misspecification. Introducing crisis and non-crisis dummy variables allows for a deeper understanding of how these funds perform in different market conditions. In both cases, neither the crisis nor the non-crisis alphas are significant, meaning that neither portfolio exhibits abnor-

Mutual funds with Sustainable Investment strategies - conventional mutual funds					
Full period			**Crisis and Non-Crisis Periods**		
Parameter	Coefficient (β)	Std. Error	Parameter	Coefficient (β)	Std. Error
α	-0.2525***	0.0378	Non-Crisis Period	-0.2635***	0.0402
	[-6.6827]			[-6.5609]	
			Crisis Period	-0.1556	0.1249
				[-1.2462]	
$M - Rf$	-0.0232***	0.0088	$M - Rf$	-0.0219**	0.0090
	[-2.6288]			[-2.4472]	
SMB	-0.3163***	0.0122	SMB	-0.3174***	0.0123
	[-25.9255]			[-25.8456]	
HML	-0.1013***	0.0118	HML	-0.1000***	0.0119
	[-8.5666]			[-8.3936]	
MOM	-0.0209**	0.0086	MOM	-0.0199**	0.0087
	[-2.4206]			[-2.2835]	

Table 4.15: Performance of an equally weighted portfolio long on the mutual funds practicing sustainable investment strategies and short on the matched conventional mutual funds with Carhart regression model. Full period regression tests: $R^2 = 0.6621$; F Test p-value = 0.0000; Ramsey RESET Test p-value = 0.6915. Crisis and non crisis period regression tests: $R^2 = 0.6627$; F Test p-value = 0.0000; Ramsey RESET Test p-value = 0.6848.

Notes: t-statistics are reported in brackets. Significance levels are denoted as follows: *** p<0.01, ** p<0.05, * p<0.10.

mal returns across different economic conditions. This suggests that both sustainable and non-sustainable funds remain highly dependent on market conditions, without generating excess returns during financial downturns. This perspective changes drastically when a portfolio which combined both groups, going long on the sustainable and short on the conventional one, is studied with Carhart regression model. Table 4.15 shows a highly statistically significant negative α through the full period and considering recessions, determining an overperformance of conventional mutual funds over sustainable ones. Furthermore, the factors' coefficients shift and become negative and still statistically significant, suggesting that the combined portfolio is directed towards big-cap, growth stocks and an anti-momentum strategy. This is maintained in the regression accounting for crisis periods. Furthermore, the t-statistic in table B.9 indicates that the means of alpha of sustainable and conventional funds differ, proving a slight underperformance of sustainable mutual funds.

4.3 Robustness Checks

The consistency of the empirical results across different sub-periods is tested. This involved dividing the sample into multiple time periods and analyzing whether the results hold consistently over these sub-periods. By conducting the analysis in distinct time intervals, the study can assess whether the observed effects are stable over time or whether they are specific to certain periods. This test of temporal stability is crucial to ensure that the conclusions drawn from the full sample are not driven by time-specific effects or changes in market conditions. Stability of the results across sub-periods will indicate the robustness of the findings and their general applicability over time. Following Nofsinger and Varma (2014), this will be tested with dummy variables accounting for four recession periods determined by National Bureau of Economic Research (NBER) (2023) to be the periods going from July 1990 to March 1991, from March 2001 to November 2001, from December 2007 to June 2009 and from February 2020 to April 2020. Overall the results did not differ when accounting for the recession period, confirming what predicted with the full period regression and the robustness of results. However, in few cases, such as the 50% cutoff point, the Momentum factor went from highly statistically significant to marginally significant 4.1 4.2, as the α in the unique portfolio regression 4.3. At the 25% cutoff point, the α of the lower-ranked portfolio gained significance 4.5, same as both portfolios at the 10% cutoff point 4.7 4.8, which was not confirmed by the unique portfolio regression 4.9 but the t-statistics for full period and for the non crisis period weren't statistically significant at 10% cutoff B.4. Furthermore, the t-statistics for crisis and non crisis period at the 25% cutoff point were statistically significant, suggesting differences in portfolio performances B.6, as for the investment screens case B.10 and the cutoff at 5% B.2.

4.4 Study Limitations

This research has a few shortcomings that must be addressed.

First of all, the data availability was restricted to lack of a time series for ESG scores for mutual funds, lack of complete data regarding mutual fund holdings and holdings' ESG scores. Moreover, ESG score time series for stocks were incomplete, leading to a further limitation on the study period that was reduced from 1965-2024 to 2007-2023.

Moreover, while the Carhart four-factor regression framework provides a robust methodology for evaluating mutual fund performance, several limitations must be considered when interpreting the results. First, the model specification may suffer from omitted variable bias, as it does not account for additional risk factors such as liquidity constraints, macroeconomic shocks, or sector exposures, which could influence fund performance. The statistical power of alpha estimates is also a concern, with high standard errors making it difficult to detect small but meaningful deviations from zero.

Furthermore, the ESG scores are not standardized, consequently by using scores from different entities the same empirical analysis can yield different results, as shown by figure 2.1. Also, greenwashing remains a concern since it is not limited by the period 2007-2023 considered, even though it can be for the period 1990-2024, but overall can't be evaluated through the use of just one entity's ESG scores.

Additionally, the matching procedure used for the analysis of sustainable investment screens is not repeated in different years, thus the study does not take in consideration a possible bias in the matching procedure.

The absence of a thorough analysis involving fund expenses, fund turnover and investment objective gives uncertainty about, for example, the comparability of α, since it is not possible to know wether a fund generates a true α or if it simply charges lower fees, or viceversa, or if the slightly negative α encountered can be a result of fund turnover (Carhart 1997), or wether the fund is sustainable or not if there is lack of ESG data and fund objective, leading to its exclusion from SRI mutual funds.

Finally, external factors affecting excess returns such as skills of mutual fund managers cannot be avoided.

Conclusion

A comparative analysis of the Carhart four-factor regression across different percentiles of mutual fund ESG performance and sample of funds that employs sustainable investment strategies, provides valuable insights into the persistence of abnormal returns. The results suggest that none of the portfolios exhibit significant positive alpha across full period, crisis, or non-crisis regressions, reinforcing the notion that mutual fund performance is largely explained by market, size, value, and momentum factors. However, lower-ranked funds (bottom 25% and bottom 10%) display marginally significant negative alpha over the full period, indicating potential long-term underperformance. One of the most striking finding is that during crisis periods, bottom 25% and bottom 10% funds exhibit significantly negative alpha, suggesting that these funds are particularly vulnerable to market downturns. Interestingly, top 10% funds also display marginally significant negative alpha in crises, implying that even highly ranked funds may not be fully insulated from downturn effects. In contrast, non-crisis period alphas for all portfolios are statistically insignificant, indicating that neither top-performing nor underperforming funds generate abnormal returns in stable market conditions. Overall, the most interesting result concerns the highly statistically significant negative α of the regression on a portfolio going long on funds applying sustainable investment screens and short on matched conventional mutual funds, suggesting a poor performance of SRI mutual funds, despite being small. These findings suggest that mutual fund returns are predominantly systematic, with little evidence of persistent manager-driven alpha.

Appendix A

Literature Review

Table A.1: Literature Review

Study	Period	No. of Funds	Results
Bauer et al. (2005)	1990-2001	55 SRI funds, 3874 non-SRI	Ethical funds have smaller size and higher expense ratio than conventional funds. The average monthly alphas of SRI funds is -0.05%. The US domestic ethical funds significantly underperform conventional domestic funds, while for US international funds the difference in returns between ethical and conventional funds is insignificant. After significant underperformance in the early 1990s, they match conventional fund performance over 1998-2001. Older ethical funds (launched before 1998) outperform younger ethical funds. All SRI funds are more growth- than value-oriented and tilted towards large capitalization stocks. (Renneboog et al. 2008b)

Continued on next page

Table A.1 – *Continued from previous page*

Study	Period	No. of Funds	Results
Bollen (2007)	1980-2002	187 SRI funds, 9189 non-SRI funds	R^2 are 81.58% for conventional and 87.12% for SRI funds. Both alpha are negative, with conventional funds alpha of -0.0025 slightly lower than SRI funds alpha of -0.0017. SRI funds in the sample are weighted toward larger capitalization stocks relative to conventional funds, consistent with the results of Bauer et al. (2005). The SRI funds also have a significantly smaller exposure to momentum stocks.
Borgers et al. (2015)	2004-2012	average of 715 mutual funds per year starting with 52 SRI funds in 2004 becoming 72 in 2012	risk-return difference is not statistically significant.
Derwall et al. (2011)	1992-2008	all stocks in KLD database	2.64% significant alpha for non-SRI stocks and 2.81% alpha not significant for SRI stocks, that became significant and higher for smaller periods.
Eccles et al. (2014)	1993-2009	90 SRI stocks and 90 conventional stocks	SRI stocks significantly outperform conventional stocks by 4.8% on a value-weighted portfolio and by 2.3% on an equal-weighted portfolio.

Continued on next page

Table A.1 – *Continued from previous page*

Study	Period	No. of Funds	Results
El Ghoul and Karoui (2017)	2003-2011	2168 mutual funds	Funds with high CSR score exhibit poor and persistent performance. The CSR score is significantly related to all factor loadings. In particular, a high fund CSR score goes hand in hand with significant exposure to small-beta stocks, large stocks, growth stocks, and contrarian stocks. These results corroborate those of Bollen (2007), who finds that relative to conventional funds, SRI funds are weighted toward larger capitalization and contrarian stocks.
El Ghoul and Karoui (2022)	2010-2017	2516 Mutual funds	Socially responsible funds outperform their less socially responsible peers.
Galema et al. (2008)	1992-2006; 1991-2004	entire database	SRI results in lower book-to- market ratios, and as a result, the alphas do not capture SRI effects. However, portfolios that score positive on diversity, environment and product have a significant impact on stock returns by lowering book-to-market ratios. SRI portfolios have lower exposures to HML factor and are growth oriented. Only community strength portfolio significantly outperforms its counterpart and employee relation score has a significant positive impact on excess returns.

Continued on next page

Table A.1 – *Continued from previous page*

Study	**Period**	**No. of Funds**	**Results**
Gil-Bazo et al. (2010)	1997-2005	86 SRI mutual funds and 1761 non-SRI mutual funds	SRI funds operated by companies specialized in SRI significantly outperform their peers, while SRI funds operated by generalist companies underperform conventional funds.
Goldreyer et al. (1999)	1981-1997	49 SRI funds and 180 non-SRI funds	The average Jensen's alpha of 29 SRI equity funds is -0.49% per annum, whereas that of 20 non-SRI equity funds is 2.78%. The difference is not significant. SRI funds using positive screens outperform the SRI funds that do not (the average monthly alphas are -0.11% and -0.81%, respectively, and the difference between them is statistically significant). (Renneboog et al. 2008b)
Hamilton et al. (1993)	1981-1985; 1986-1990	32 SRI funds and 320 non-SRI funds	For 17 SRI funds established before 1985, the average alpha is -0.06% per month, which is higher than the average monthly alpha (-0.14%) of 170 non-SRI funds (the difference is not significant). Meanwhile for the 15 SRI funds with shorter history, i.e. established after 1985, the average alpha is -0.28% per month, which is worse than the average monthly alpha (-0.04%) of the corresponding 150 non-SRI funds. (Renneboog et al. 2008b)
Kempf and Osthoff (2007)	1992-2004	all stocks in S&P500 and DS400	portfolios formed with positive screening approach or best-in-class screening approach earn high abnormal returns (8.7% per year).

Continued on next page

Table A.1 – *Continued from previous page*

Study	Period	No. of Funds	Results
Nofsinger and Varma (2014)	2000-2011	240 SRI funds and 720 non-SRI funds	Overall, the alphas for SRI and conventional funds are insignificantly negative and not different from each other. However, in non-crisis periods, conventional funds outperform SRI funds by an annualized 0.67–0.95%, depending on the factor model used. But in crisis periods, SRI funds outperform by 1.61–1.70%.
Renneboog et al. (2008a)	1991-2003	440 SRI funds and 16036 non-SRI funds	SRI funds strongly underperform domestic benchmark portfolios (such as the Fama-French-Carhart factors). In particular, the average risk-adjusted returns of the SRI funds range from -2.2% to -6.5% per annum. However, comparing the alphas of the SRI funds with those of matched conventional funds, there is no statistically significant evidence that SRI funds underperform their conventional counterparts. (Renneboog et al. 2008b)

Continued on next page

Table A.1 – *Continued from previous page*

Study	**Period**	**No. of Funds**	**Results**
Renneboog et al. (2011)	1992-2003	321 SRI funds and 3113 non-SRI funds	While SRI funds with Sin/Ethical screens or Environmental screens significantly underperform matched conventional funds, some SRI attributes have a positive impact on future returns. In particular, funds with a policy of activism can expect 4% higher returns per annum on a risk-adjusted basis. Finally, funds that receive more inflows neither outperform nor underperform their benchmarks or conventional funds.
Schroder (2004)	1990-2002	46 funds	The monthly alphas range from -2.06% to 0.87%. 38 out of the 46 alphas are negative; only 4 of them are significant at 0.05 level. SRI funds do not significantly underperform the benchmark portfolio consisting of both large stocks and small stocks. (Renneboog et al. 2008b)
Statman (2000)	1990-1998	31 SRI funds and 62 non-SRI funds	The average monthly alpha is -0.42% for SRI funds and -0.62% for non-SRI funds; the difference is not significant (t-statistics = 1.84). The DSI 400 index has a higher Sharpe ratio than the S&P 500 index (0.97 vs. 0.92). (Renneboog et al. 2008b)

Appendix B

Individual regressions statistics

The following tables present the results of Carhart four-factor regression analyses performed on two sets of mutual funds: Sustainable and Responsible Investment (SRI) funds and conventional mutual funds. The regression model includes the following factors: Alpha (α), Market Risk Premium (M-Rf), Size factor (SMB), Value factor (HML), and Momentum (MOM). For each of these factor, after all the individual regressions have been computed, their means, standard deviation and number of times a pvalue was under 1%, 5% and 10%, are registered. Moreover, the sample size is displayed in the lower part of the tables, along with a t-statistic of the difference of the average α between the samples.

Table B.1: Average performance of top and bottom 5% ESG-ranked mutual funds with Carhart regression model

	SRI Mutual Funds					**Conventional Mutual Funds**				
Parameter	**Mean**	**Std**	**p1%**	**p5%**	**p10%**	**Mean**	**Std**	**p1%**	**p5%**	**p10%**
α	-0.0252	1.6881	1	27	32	-0.0838	0.6224	8	27	17
$M-Rf$	0.9232	0.3387	785	12	0	0.9818	0.2796	436	17	5
SMB	-0.0182	0.2721	68	131	86	0.6578	0.4029	289	43	31
HML	0.3658	6.8514	194	89	43	-0.0416	0.3873	144	85	23
MOM	-0.0731	0.2079	37	97	102	0.0828	0.2500	41	36	72
Tot.			807					481		
α t-statistic					0.8899					

Table B.2: Average performance of top and bottom 5% ESG-ranked mutual funds with Carhart regression model introducing crisis periods dummy variables

	SRI Mutual Funds					**Conventional Mutual Funds**				
Parameter	**Mean**	**Std**	**p1%**	**p5%**	**p10%**	**Mean**	**Std**	**p1%**	**p5%**	**p10%**
Non-Crisis	-0.0330	0.3321	0	6	7	0.1000	0.7584	0	7	3
Crisis	-0.6130	0.9968	3	11	26	-0.1688	1.3308	3	6	9
$M - Rf$	0.9428	0.2438	241	1	0	0.9822	0.2798	178	1	1
SMB	-0.0240	0.3035	17	37	22	0.6594	0.4022	120	10	9
HML	0.1012	1.2594	55	21	13	-0.0477	0.3895	69	20	4
MOM	-0.0792	0.2118	16	28	14	0.0793	0.2541	22	47	16
Tot.			807					481		
α **t-statistic**	Non-Crisis				-3.644					
α **t-statistic**	Crisis				-6.337					

Table B.3: Average performance of top and bottom 10% ESG-ranked mutual funds with Carhart regression model

	SRI Mutual Funds					**Conventional Mutual Funds**				
Parameter	**Mean**	**Std**	**p1%**	**p5%**	**p10%**	**Mean**	**Std**	**p1%**	**p5%**	**p10%**
α	-0.0173	1.5745	15	59	66	-0.0654	0.9706	16	53	35
$M - Rf$	0.9249	0.4646	1309	17	2	0.9695	0.4475	824	18	4
SMB	-0.0160	0.2807	221	157	95	0.5413	0.4607	517	103	37
HML	0.2688	5.3227	468	164	90	0.0300	0.4427	291	122	44
MOM	-0.0407	0.1951	93	179	126	-0.0077	0.4649	72	67	97
Tot.			1340					879		
α **t-statistic**					0.4834					

Table B.4: Average performance of top and bottom 10% ESG-ranked mutual funds with Carhart regression model introducing crisis periods dummy variables

	SRI Mutual Funds					**Conventional Mutual Funds**				
Parameter	**Mean**	**Std**	**p1%**	**p5%**	**p10%**	**Mean**	**Std**	**p1%**	**p5%**	**p10%**
Non-Crisis	-0.0113	0.3156	5	13	19	-0.0057	0.6258	2	24	23
Crisis	-0.6144	1.2365	17	72	43	-0.5122	1.3716	13	27	24
$M - Rf$	0.9516	0.2806	486	1	0	0.9240	0.9163	367	0	1
SMB	-0.0202	0.3036	111	51	40	0.4475	1.9740	259	39	22
HML	0.1100	1.0020	175	62	51	0.1081	1.3820	158	37	51
MOM	-0.0456	0.1964	38	47	46	0.0073	0.4091	43	50	35
α **t-statistic**	Non-Crisis				-0.246					
α **t-statistic**	Crisis				-1.784					

Table B.5: Average performance of top and bottom 25% ESG-ranked mutual funds with Carhart regression model

	SRI Mutual Funds					Conventional Mutual Funds				
Parameter	**Mean**	**Std**	**p1%**	**p5%**	**p10%**	**Mean**	**Std**	**p1%**	**p5%**	**p10%**
α	-0.1721	4.1419	44	166	164	-0.1071	0.8272	49	138	121
$M - Rf$	0.9765	0.9620	2215	33	5	1.0019	0.2549	1934	16	7
SMB	0.0026	0.3714	581	260	185	0.3317	0.4665	1084	187	119
HML	0.1856	4.0646	1134	184	70	-0.0163	0.8015	897	218	104
MOM	-0.0353	0.3460	248	295	200	-0.0053	0.3007	221	205	261
Tot.			2272					1998		
α **t-statistic**					-0.7316					

Table B.6: Average performance of top and bottom 25% ESG-ranked mutual funds with Carhart regression model introducing crisis periods dummy variables

	SRI Mutual Funds					Conventional Mutual Funds				
Parameter	**Mean**	**Std**	**p1%**	**p5%**	**p10%**	**Mean**	**Std**	**p1%**	**p5%**	**p10%**
Non-Crisis	-0.0437	0.2633	25	39	70	-0.0002	0.4642	10	40	42
Crisis	-0.4366	1.1395	60	144	123	-0.4854	1.4074	40	62	61
$M - Rf$	0.9737	0.9700	1166	0	1	0.9838	0.5164	914	4	0
SMB	0.0183	0.6887	441	143	91	0.2868	1.3255	674	75	35
HML	0.0911	0.6299	664	107	48	0.0123	1.1751	491	120	46
MOM	-0.0270	0.6118	181	133	99	-0.0031	0.2793	119	121	80
Tot.			2272					1998		
α **t-statistic**	Non-Crisis					-3.698				
α **t-statistic**	Crisis					1.234				

Table B.7: Average performance of top and bottom 50% ESG-ranked mutual funds with Carhart regression model

	SRI Mutual Funds					Conventional Mutual Funds				
Parameter	**Mean**	**Std**	**p1%**	**p5%**	**p10%**	**Mean**	**Std**	**p1%**	**p5%**	**p10%**
α	-0.0492	1.9224	98	205	218	-0.0766	0.9874	74	225	174
$M - Rf$	0.9624	0.5161	2947	13	16	0.9796	0.3096	3070	12	3
SMB	0.0363	0.4987	1204	343	215	0.2015	0.4511	1649	271	184
HML	0.0772	0.3280	1803	270	112	0.0145	1.0042	1761	259	132
MOM	-0.0368	0.2252	428	353	243	-0.0193	0.2473	437	339	254
Tot.			3006					3131		
α **t-statistic**					0.6980					

Table B.8: Average performance of top and bottom 50% ESG-ranked mutual funds with Carhart regression model introducing crisis periods dummy variables

	SRI Mutual Funds					Conventional Mutual Funds				
Parameter	**Mean**	**Std**	**p1%**	**p5%**	**p10%**	**Mean**	**Std**	**p1%**	**p5%**	**p10%**
Non-Crisis	-0.0431	0.2370	62	150	101	-0.0377	0.3206	19	81	82
Crisis	-0.3688	1.0871	165	180	175	-0.3502	1.2417	119	115	110
$M-Rf$	0.9826	0.9068	2028	0	0	0.9822	0.5632	1824	3	0
SMB	0.0529	0.6926	908	230	125	0.1760	1.0842	1211	144	81
HML	0.0738	0.3114	1337	135	97	0.0284	1.2153	1145	162	82
MOM	-0.0330	0.5173	303	263	139	-0.0212	0.2487	284	246	103
Tot.			3006					3131		
α **t-statistic**	Non-Crisis					-0.752				
α **t-statistic**	Crisis					-0.625				

Table B.9: Average performance of mutual funds applying sustainable investment techniques and conventional mutual funds with Carhart regression model

	SRI Mutual Funds					Conventional Mutual Funds				
Parameter	**Mean**	**Std**	**p1%**	**p5%**	**p10%**	**Mean**	**Std**	**p1%**	**p5%**	**p10%**
α	-0.1434	0.1907	12	14	18	-0.0919	0.1736	38	85	79
$M-Rf$	0.9640	0.0778	149	0	0	1.0185	0.1011	873	0	0
SMB	0.0338	0.1925	66	18	11	0.3972	0.3535	729	29	25
HML	0.0391	0.1912	78	17	7	0.0357	0.3304	661	59	32
MOM	-0.0063	0.0883	28	10	11	0.0246	0.1166	276	104	88
Tot.			149					880		
α **t-statistic**					-3.087					

Table B.10: Average performance of mutual funds applying sustainable investment techniques and conventional mutual funds with Carhart regression model introducing crisis periods dummy variables

	SRI Mutual Funds					Conventional Mutual Funds				
Parameter	**Mean**	**Std**	**p1%**	**p5%**	**p10%**	**Mean**	**Std**	**p1%**	**p5%**	**p10%**
Non-Crisis	-0.1062	0.1350	6	16	6	-0.0788	0.1435	29	59	63
Crisis	-0.1390	0.7395	3	8	8	-0.2835	0.9435	43	73	61
$M-Rf$	0.9644	0.0774	114	0	0	1.0176	0.1012	841	0	0
SMB	0.0334	0.1927	58	14	5	0.3995	0.5632	715	22	27
HML	0.0386	0.1903	61	16	7	0.0348	0.3271	646	53	35
MOM	-0.0057	0.0875	23	11	10	0.0250	0.1160	276	105	76
Tot.			149					880		
α **t-statistic**	Non-Crisis					-2.270				
α **t-statistic**	Crisis					2.112				

Bibliography

Bauer, R., Koedijk, K., & Otten, R. (2005). International evidence on ethical mutual fund performance and investment style. *Journal of Banking and Finance*, *29*(7), 1751–1767.

Berg, F., Koelbel, J., & Rigobon, R. (2019). Aggregate confusion: The divergence of esg ratings. *Forthcoming Review of Finance*.

Bollen, N. P. (2007). Mutual fund attributes and investor behavior. *Journal of Financial and Quantitative analysis*, *42*(3), 683–708. https://doi.org/10.1017/s0022109000004142

Borgers, A., Derwall, J., Koedijk, K., & Ter Horst, J. (2015). Do social factors influence investment behavior and performance? evidence from mutual fund holdings. *Journal of Banking and Finance*, *60*, 112–126.

Brown, S., Goetzmann, W., Ibbotson, R., & Ross, S. (1992). Survivorship bias in performance studies. *Review of Financial Studies*, *5*, 553–580.

Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance*, *52*, 57–82.

Derwall, J., Koedijk, K., & Ter Horst, J. (2011). A tale of values-driven and profit-seeking social investors. *Journal of Banking and Finance*, *35*, 2137–2147.

Eccles, R. G., Ioannou, I., & Serafeim, G. (2014). The impact of corporate sustainability on organizational processes and performance. *Management Science*, *60*(11), 2835–2857.

El Ghoul, S., & Karoui, A. (2017). Does corporate social responsibility affect mutual fund performance and flows? *Journal of Banking and Finance*, *77*, 53–63.

El Ghoul, S., & Karoui, A. (2022). Fund performance and social responsibility: New evidence using social active share and social tracking error. *Journal of Banking and Finance*, *143*(106598).

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.

Galema, R., Plantinga, A., & Scholtens, B. (2008). The stocks at stake: Return and risk in socially responsible investment. *Journal of Banking and Finance*, *32*(12), 2646–2654.

Gil-Bazo, J., Ruiz-Verdu, P., & Santos, A. A. P. (2010). The performance of socially responsible mutual funds: The role of fees and management companies. *Journal of Business Ethics*, *94*(2), 243–263.

Goldreyer, E., Ahmed, P., & Diltz, J. D. (1999). The performance of socially responsible mutual funds: Incorporating sociopolitical information in portfolio selection. *Managerial Finance*, *25*(1), 23–36.

Hamilton, S., Jo, H., & Statman, M. (1993). Doing well while doing good? the investment performance of socially responsible mutual funds. *Financial Analysts Journal*, *49*(6), 62–66.

IPCC. (2023a). *Climate change 2023: Synthesis report. contribution of working groups I, II and III to the sixth assessment report of the intergovernmental panel on climate change [core writing team, h. lee and j. romero (eds.)]* (IPCC report). IPCC.

IPCC. (2023b). Climate change 2023: Synthesis report. contribution of working groups i, ii and iii to the sixth assessment report of the ipcc.

Jensen, M. (1968). The performance of mutual funds in the period 1945–1964. *The Journal of Finance*, *23*(2), 389–416.

Kempf, A., & Osthoff, P. (2007). The effect of socially responsible investing on portfolio performance. *European Financial Management*, *13*(5), 908–922.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics*, *47*(1), 13–37.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, *7*(1), 77–91.

Mooney, B., Tauschinski. (2025). World breaches 1.5c global warming target for first time in 2024. *Financial Times*. https://www.ft.com/content/fd914266-71bf-4317-9fdc-44b55acb52f6

Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica*, *34*(4), 768–783.

National Bureau of Economic Research (NBER). (2023). Business cycles data. https://www.nber.org/research/data/us-business-cycle-expansions-and-contractions

Nofsinger, J., & Varma, A. (2014). Socially responsible funds and market crises. *Journal of Banking and Finance*, *48*, 180–193.

Rayner, K., & Rosanna, C. (2022). *Global sustainable investment review* (GSIA review). Global Sustainable Investment Alliance (GSIA). https://www.gsi-alliance.org/wp-content/uploads/2023/12/GSIA-Report-2022.pdf

Refinitiv LSEG. (2024). *Environmental, Social, and Governance Scores Guide*. London Stock Exchange Group (LSEG). https://www.lseg.com/content/dam/data-analytics/en_us/documents/methodology/lseg-esg-scores-methodology.pdf

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.

Renneboog, L., Ter Horst, J., & Zhang, C. (2008a). 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. (2008b). Socially responsible investments: Institutional aspects, performance, and investor behavior. *Journal of Banking and Finance*, *32*, 1723–1742.

Schroder, M. (2004). The performance of socially responsible investments: Investment funds and indices. *Financial Markets and Portfolio Management*, *18*(2), 122–142.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, *19*(3), 425–442.

Statman, M. (2000). Socially responsible mutual funds. *Financial Analysts Journal*, *56*(3), 30–39.

UNEP. (2024). *Emissions gap report 2024: No more hot air . . . please! with a massive gap between rhetoric and reality, countries draft new climate commitments* (UNEP report). UNEP. Nairobi. 11822/46404.