Ari Kauppi

ESG-Stability Factor Performance

Constructing and Testing ESG Stability and Momentum in Sector-Neutral Portfolios

School of Accounting and Finance

Master’s Degree in Finance

Master’s Degree Programme in Finance

Vaasa 2025

VAASAN YLIOPISTO
School of Accounting and Finance
Author: Ari Kauppi
Thesis title: ESG-Stability Factor Performance : Constructing and Testing ESG Stability and Momentum in Sector-Neutral Portfolios
Degree: Master’s degree
Supervisor: Klaus Grobys
Year of graduation: 2025 **Number of pages:** 95

ABSTRACT:
This thesis tests whether the time-series behaviour of firms’ environmental, social, and governance (ESG) scores contains priced information in U.S. large-cap equities. Using a monthly panel of S&P 500 constituents, I form sector-neutral, value-weighted long–short portfolios based on two signals: ESG momentum (recent improvement in an ESG score) and ESG stability (low variability of that score over time). To reduce mechanical tilts, scores are aligned for a risk-rating polarity change, winsorised cross-sectionally, and rank-neutralised by size and sector. Performance is evaluated with standard asset-pricing tests—Fama–French five-factor regressions and models augmented with price momentum and a defensive/low-volatility factor—using Newey–West inference. Model adequacy is further assessed with Gibbons–Ross–Shanken (GRS) tests.
The evidence points to a robust premium for ESG stability: portfolios long stable scorers and short unstable scorers earn positive, statistically significant abnormal returns that are not explained by standard factors. By contrast, ESG momentum does not deliver positive abnormal performance; at intermediate horizons it underperforms. GRS tests indicate that adding the stability factor improves the joint pricing of test portfolios relative to FF5 alone.
Overall, the results suggest that firms with smoother, less noisy ESG trajectories outperform those with volatile trajectories, while recent ESG score improvements do not translate into excess returns. The findings are robust across reasonable window choices and weighting schemes and highlight stability as an incremental source of cross-sectional variation in returns.

Keywords: ESG, responsible investing, ESG investing, socially responsible investing

Contents

List of Figures 5

List of Tables 5

Abbreviations 7

1 Introduction 8

1.1 Background 9

1.2 Research questions and hypothesis 12

2 Socially Responsible Investing 15

2.1 History of Socially Responsible Investing 15

2.2 Responsible Investment Strategies 17

2.2.1 Shareholder advocacy 18

2.2.2 Community Investment 18

2.3 ESG Investment Strategies 18

2.3.1 Screening 18

2.3.2 ESG integration 19

2.3.3 Impact investing 19

2.3.4 ESG Momentum 20

2.3.5 ESG Stability 23

3 Theoretical framework 25

3.1 Efficient Market Hypothesis 25

3.1.1 Behavioral Finance and ESG Momentum 27

3.1.2 Tests of the Efficient Market Hypothesis 27

3.2 Market anomalies 29

3.2.1 The Small firm Effect 30

3.2.2 Momentum anomaly and ESG 30

3.3 Portfolio evaluation metrics 31

3.3.1 CAPM 31

3.3.2 Fama-French three factor model 36
3.3.3 Carhart four-factor model 38
3.3.4 Five-factor and six-factor models 39

4 Previous studies on ESG momentum and stability 41
4.1 Nagy, Cogan, and Sinnreich (2013) 41
4.2 Nagy, Kassam & Lee (2016) 42
4.3 Verheyden, Eccles, Feiner (2016) 42
4.4 Giese, Lee, Melas & Nishikawa (2019) 43
4.5 Chen & Yang (2020) 45
4.6 Magnani, Guidolin & Berk (2024) 46

5 Data and methodology 48
5.1 Data 48
5.1.1 ESG data 48
5.1.2 Defining a peer group 51
5.1.3 Sustainalytics' ESG Data 54
5.1.4 LSEG data 56
5.2 Methodology 57
5.2.1 Portfolio construction 57
5.2.2 ESG score neutralization 58
5.2.3 Portfolio construction 59

6 Results 62
6.1 Descriptive statistics 62
6.2 Portfolio analysis and factor models 64
6.2.1 Robustness testing 68
6.3 Overall Discussion 69
6.4 Limitations and Further Research 70

Bibliography 72

Appendices 84

Appendix 1. TISFD follows in the footseps of environmental taskforces 84
Appendix 1. Summary of different ESG investment styles 85
Appendix 4. Large cap and geographic biases in ESG ratings 86
Appendix 2. Summary of different ESG strategies 87
Appendix 8. Factor regressions in full 88

List of Figures

Figure 1 Overview of SRI Strategies in Europe (2013-2018) (Eurosif 2016; 2018) 17
Figure 2 Eurosif's methodology to classify sustainability-related investments (2021b) 21
Figure 3 Overview of papers which find positive alpha from ESG strategies (Bruno, Esakia, & Goltz, 2022, 1) 23
Figure 4 CAPM investment opportunities (E. F. Fama & French, 2006, 27) 34
Figure 5 Entries and exits in S&P 500 (SPY point-in-time constituent data from LSEG) 49
Figure 6 MSCI vs. Sustainalytics rankings at the beginning of 2019 (Dimson, Marsh, & Staunton, 2020, 3) 53
Figure 7 Historic Correlation of MSCI's and LSEG's ESG Ratings 2012-2018 (LaBella, Sullivan, Russell, & Novikov, 2019, 3) 54
Figure 8 Sustainalytics Risk Framework (Sustainalytics, 2025a) 55
Figure 9 Top drawdowns 64
Figure 10 Drawdowns from prior peak (path) 65
Figure 11 Cumulative wealth of different strategies and S&P 500 (2016-2024) 68
Figure 12 TISFD follows in the footsteps of environmental task forces (Thomson & Bastit, 2024) 84
Figure 13 Large cap bias present in ESG ratings(LaBella et al., 2019, 5) 86

List of Tables

Table 1 ESG Incorporation strategies (adapted from UN PRI 2023; 2025) 20

Table 2 Various different units used to measure Employee Health and Safety in random sample of 50 large, publicly listed companies (Kotsantonis & Serafeim, 2019, 52) 50

Table 3 Summary statistics and ESG scores for five size-sorted quantiles (adapted from Magnani, Guidolin, & Berk, 2024, 672) 63

Table 4 Summary statistics of returns and ESG scores for industries 63

Table 5 ESG Momentum (long–short) results across look-back horizons H 66

Table 6 ESG Stability (VOL) long–short results across volatility windows W 66

Table 7 ESG volatility portfolio regression with $W = 24$ 67

Table 8 GRS tests across factor sets (test portfolios: 5 quintiles; monthly data) 68

Table 9 Summary of CSR investment strategies models (Global Sustainable Investment Alliance, 2013) 85

Table 10 Overview of Papers on ESG and Positive Alpha 87

Abbreviations

CSR Corporate Social Responsibility

CSP See CSR

SRI Sustainable and Responsible Investing

PRI Principles of Responsible Investing. A set of guidelines published by the UN in 2006.

EMH Efficient Market Hypothesis

GSIA Global Sustainable Investment Alliance

ESG Environmental, Social, and Governance

Eurosif The European Social Investment Forum

CFP Corporate Financial Performance

RI Responsible Investing

CAPM Capital Asset Pricing Model

MPT Modern Portfolio Theory

MKT Market Risk Premium

WML Winners Minus Losers

CFP Corporate Financial Performance

SFDR Sustainable Finance Disclosure Regulation (EU)

FMP Financial Market Participants

US SIF (US) Sustainable Investment Forum, member of GSIA

SDSN (UN) Sustainable Development Solutions Network

TRBC The Thomson Reuters Business Classification

1 Introduction

Sustainable and Responsible Investing (SRI) has evolved significantly over the past few decades, transitioning from a niche concept to a mainstream strategy embraced by a diverse group of investors. SRI lacks clear definition but its central principle is the integration of ethical, environmental, and social considerations into investment decisions. This shift in investor priorities mirrors broader societal calls for corporate accountability, transparency, and sustainable business practices. It is also reflected by the increased volume of SRI investments. In the US, between 1995-2008 the volume grew from 0.64 to 12 trillion USD and in Europe more than 60 % of professionally managed assets incorporated some form of SRI in practices in 2022 (EUROSIF, 2022; US SIF, 2021). As a result, investors, the general public, and stakeholders increasingly demand businesses to address these pressing global issues as well as comprehensive information on how companies address their social and environmental responsibilities. Reflecting this changing landscape, the financial sector has begun to move beyond its traditional sole focus on monetary returns, giving rise to a dynamic SRI industry.

To meet this growing demand for CSR-related information, a wide array of ratings agencies have begun providing detailed assessments of companies' adherence to Corporate Social Responsibility (CSR) principles, offering insights into their social and environmental performance. However, the proliferation of ESG ratings providers has also led to lack of consistency, terminology, and different interpretations of ESG materiality for different businesses (OICU-ISCO, 2021, 2,60,62). In 2019, the size of the global ESG ratings market was estimated to be $200m and expected to grow to $500m within five years (Nauman, 2019).

As companies integrate new environmental, social, and governance (ESG) considerations into their day-to-day operations, the effect of these factors on financial performance has become a major focus of research. Central to this inquiry is whether incorporating the non-financial criteria of ESG and socially responsible investing (SRI) hinders or enhances returns. Some recent studies have suggested that including these risk factors may lead to abnormal returns. For instance, van Duuren, Plantinga, and Scholtens (2016) find that integrating ESG considerations can positively influence performance. However, the evidence remains unclear and several studies also point to the contrary.

The picture is further complicated by the the fact that key terms such as sustainability, corporate social responsibility, and socially responsible investing often remain loosely defined, leading to inconsistent usage, occasional confusion, and even opportunistic greenwashing (see, e.g. IOSCO, 2021). For investors, results may also vary depending on the SRI strategies chosen. The most popular form of SRI investing has been positive/negative screening, where particular stocks or industries may be excluded or included due to ethical concerns, for example. More recently, ESG integration has become more relevant due to the aforementioned rise of ESG rating providers and wider availability of ESG data (EUROSIF, 2018).

This study aims to contribute to this discussion by examining the relationship between ESG factors and financial performance. More specifically, this study contributes to a fairly recent SRI strategy called ESG momentum and ESG stability. The following section 1.1 gives a short introduction into the background of the study and the purposes and research questions of this study will be given in section 1.2. This strategy is tested using the S&P 500 companies. More reasons on why this was chosen as the dataset can be found in section 5.1.3.

1.1 Background

Environmental, Social, and Governance (ESG) metrics have become a cornerstone of evaluating corporate responsibility and sustainability. These metrics assess how companies perform across three broad dimensions: environmental stewardship, social impact, and governance practices. However, the precise composition and weight of ESG criteria as well as materiality considerations of different ESG issues may vary significantly across different rating agencies (e.g. EUROSIF, 2018; IOSCO, 2021). Prominent agencies such as Refinitiv, Sustainalytics, MSCI, and S&P Global provide ESG scores that guide investors in assessing corporate behavior. According to Puttonen and Puttonen (2021), these scores reflect a weighted evaluation of diverse factors, influenced by industry characteristics and agency-specific methodologies. ESG scores are particularly crucial for publicly listed companies, as they serve as indicators of their commitment to sustainable practices.

The relevance of ESG metrics is closely related to the objectives of SRI and CSR. For instance, The European Social Investment Forum (Eurosif) (2022) defines SRI as a long-term investment strategy that incorporates ESG factors in security selection and portfolio management. SRI emphasizes not only financial returns but also the broader societal and environmental impact of investments. Some authors, such as Brzeszczyński and McIntosh (2014) argue that investors implementing an SRI strategy may be willing to a portion of their financial gains to better achieve their environmental, social, and/or ethical agenda.

ESG metrics have become an important addition for investors implementing Sustainable and Responsible Investing (SRI) strategies. Historically, these strategies often relied on screening or excluding companies based on their involvement in certain industries or social and environmental issues. However, this approach can overlook firms that—though not operating in a screened industry—still perform poorly in areas such as social responsibility or governance.

By incorporating ESG metrics, investors can go beyond simple industry exclusions and evaluate companies based on measurable criteria of environmental, social, and governance performance. While setting specific ESG thresholds inevitably involves some arbitrariness and may only be appropriate for a particular universe of assets, the growing reliance on these metrics has fueled extensive research on the relationship between social responsibility and Corporate Financial Performance (CFP). Although the findings vary, a slight majority of studies suggest a positive correlation between robust ESG practices and financial returns.

A seminal meta-analysis by Friede, Busch, and Bassen (2015), based on over 2,000 studies, found that 90% reported a non-negative relationship between ESG and corporate financial performance (CFP), with most indicating a positive link. More spefically, 50% of the studies covered by Friede et al's analysis found a positive relationship between ESG and CFP, 40% found mixed results, while around 10% found a negative relationship between ESG and CFP. Similarly, von Wallis and Klein (2015) and Revelli and Viviani (2015) each conducted meta-analyses on the financial effects of socially responsible investing (SRI). The former reported a modestly positive relationship, while the latter concluded that SRI does not harm returns—though it could not confirm a clear advantage. Although these findings do not strongly reinforce the core arguments of this thesis, they suggest that adopting ESG standards is unlikely to impose significant costs or deliver substantial benefits.

These results appear to conflict with Markowitz's (1952) Modern Portfolio Theory (MPT), reviewed in Section 3.3.1, which assumes investors are homogeneous and rational, basing their decisions solely on risk and return. By this logic, restricting the investment universe for any reason beyond risk and return should lead to poorer performance. However, the findings outlined above seem to indicate otherwise.

Using ESG ratings to inform investment decisions may not lead to optimal performance either. Some studies suggest a more nuanced relationship between ESG and financial performance. Bannier, Bofinger, and Rock (2019) argue that investing solely in companies with the highest ESG ratings does not guarantee superior returns. Their analysis demonstrates that portfolios long on stocks with high ESG ratings and short on those with low ratings often yield significantly negative abnormal returns. This observation suggests that high ESG ratings may already be priced into the market, reducing the potential for excess returns. Furthermore, the authors note that low ESG-rated stocks often offer higher returns, compensating for their associated risks. This dynamic indicates that investors seeking optimal financial performance might need to consider additional metrics beyond ESG ratings.

Studies such as Nagy, Kassam, and Lee (2016) and Nagy, Cogan, and Sinnreich (2013), Verheyden, Eccles, and Feiner (2016), Giese, Nagy, and Lee (2020) seem to confirm Bannier's findings that focusing on companies with improving ESG ratings can yield positive abnormal returns. For example, Nagy et al. (2016) found that portfolios oriented towards companies with upward ESG rating trends significantly outperformed those focused on high ESG-rated companies. Other studies find similar but often contradicting relationships between ESG scores and alpha. For example, Magnani, Guidolin, and Berk (2024) find that trading on ESG stability over momentum may provide better results, hinting that investors seem to reward predictability in ESG behavior more consistently than rapid improvements. These findings suggest that tracking ESG momentum may offer a promising approach for investors aiming to achieve both financial and ethical objectives.

1.2 Research questions and hypothesis

Building on this literature, the present study investigates whether an ESG momentum strategy can capture abnormal returns while also rewarding companies that progressively enhance their social and environmental performance. Especially, this thesis focuses on trying to replicate the results of Magnani, Guidolin and Berk (2024). These results suggest that two ESG signals – ESG momentum (recent improvement in a firm's ESG score) and ESG stability (low volatility of a firm's ESG score) – could deliver economically and statistically abnormal returns within the S&P 500. Following Magnani et. al's lead, I form sector-neutral, value-weighted portrolios from ESG rank-neutralised ESG scores. Therefore, the research questions of this thesis are as follows:

1. Does an S&P 500 long-short protfolio that is long high-ESG-momentum firms and short low-ESG-momentum firms earn positive risk-adjusted returns?
2. Does an S&P 500 long-short portfolio that is long stable (low-volatility) ESG firms and short unstable ESG firms earn postiive risk-adjusted returns?
3. Are any such returns distinct from standard equity factors, such as Fama-French 5 and in robustness the price-momentum factor and a defensive / low-volatility factor?

Consequently, the null hypothesis of this thesis are as follows. The first hypothesis evaluates whether an ESG-momentum strategy earns any return that cannot already be explained by standard equity factors (Fama-French 5-factors). Letting r_i^{MOM} denote the return on a long-short portfolio that is long on firms with the highest recent improvement in rank-neutralised ESG scores and short firms with the lowest improvement (constructed sector-neutrally, value-weighted with lagged market capitalisation, and applying a one-month skip), is used to estimate a time-series model:

$$\alpha^{\text{MOM}} = 0 \text{ in } r_t^{\text{MOM}} = \alpha + \beta' f_t^{\text{FF5}} + \epsilon_t \tag{1}$$

where f_t^{FF5} are the Fama-French five factors (Mkt-RF, SMB, HML, RMW, CMA).

$$H_0 - \text{MOM} : \alpha^{\text{MOM}} = 0 \tag{2}$$

This null hypothesis states that that once the factor exposures are accounted for, the ESG-momentum spread should have no abnormal performance and that any average return is just compensation for its loading on the FF5 factors.

$$H_1 - \text{MOM} : \alpha^{\text{MOM}} > 0$$

The alternative hypothesis H_1-MOM is one-sided, because prior evidence from Magnani et al. indicates a positive premium. Rejecting $H_0 -$ MOM would therefore imply that firms with improving ESG profiles deliver excess returns not captured by FF5, i.e., ESG-momentum would have incremental pricing power. Failing the which if this thesis fails to reject, would mean that ESG momentum has no incerememental pricing power in the S&P 500 beyond standard factors. The alternative hypothesis H1-MOM is that $\alpha^{\text{MOM}} > 0$ and that this is one-sided, because Magnani et al. report a positive premium. Similar tests are also carried out for testing ESG stability as a factor with similar assumptions. However, if ESG stability is found to indicate a positive premium, the conclusion would be that ESG stability would have incremental pricing power.

The remainder of this thesis is organized as follows. Section 2 reviews existing literature on SRI and ESG investing, clarifying foundational concepts and debating their varied interpretations. Section 3 outlines the theoretical framework underpinning ESG and momentum strategies. Next, Section 5 describes the data sources, sample selection, and empirical methodology employed in the study. Section 6 presents and discusses the findings, with particular attention to the performance and feasibility of ESG momentum and ESG stability strategies. Finally, Section 6.3 concludes the thesis by summarizing key insights, acknowledging limitations, and suggesting avenues for future research.

2 Socially Responsible Investing

This section introduces SRI and other related concepts, which are required to understand SRI as a phenomenon. The first section introduces the history and current state of SRI. The second section 2.2 introduces some of the most common and fundamental SRI strategies implemented by investors. Appendix 6.4 includes a short summary of different ESG investment strategies, where the different styles are described in more detail than here.

2.1 History of Socially Responsible Investing

Several authors (see, e.g., Renneboog, Ter Horst, and Zhang, 2008) trace back the roots of Socially Responsible Investing to its ancient origins in the religious traditions of Judaism and its rules regarding the ethical use of wealth. Later, this tradition was continued by Methodists and Quakers, who forbade investing in the so-called ”sin stocks” of companies involved in tobacco, gambling, alcohol or slavery industries (Martini, 2021; Renneboog et al., 2008; Sherwood & Pollard, 2018). This could be seen as a front-runner in screening investments (for more details about screening, see 2.3.1). In more recent research, sin stocks have been used to study the impact of social norms on the market. Enterprises involved in these activities seem to be exposed to significant ”price effect” caused by them being shunned by institutional investors. This in turn increases their cost of capital, affecting their stock prices and eventually the returns of sin stocks (Hong & Kacperczyk, 2009), underlining the importance of corporate social responsibility at least within the modern financial sector.

Later SRI's origins can be linked to the anti-war, and anti-racist, and various other social movements of the 1960s. The Pax World Fund, which Renneboog et al. (2008) see as the first modern SRI fund, was founded in 1971 and catered to investors by not investing in weapon contractors.

For example, Sherwood and Pollard, 12 see the modern era as beginning in the 1990s, heralded by the creation of the Domini 400 Social Index (now MSCI KLD 400 Social Index). This was further helped by Micheal Jantzi's development of the Jantzi Model to help measure the See CSR (CSP), which consisted of the following seven pillars:

1. Community issues
2. Diverse workplace
3. Employee relations
4. Environmental performance
5. International
6. Product and business practice
7. Other, which included, for example, compensation, proxy voting, and ownership in other companies

According to Sherwood and Pollard (2018, 13), these seven pillars became established as a foundational qualitative way to measure the sustainability of corporate practices. Later, with the development of the Jantzi's model, which assigned companies scores between $[-2, 2]$ in each of the pillars signifying either severe concern or indicating major success within a pillar (Fauzi, 2009).

Sherwood and Pollard (2018, 12) see Jantzi's model as serving as the beginnings of the ESG rating system, simultaneously providing cross-sectional data, which aided in predicting the return distributions of assets in different kinds of asset class. Jantzi's model also allowed companies to measure to which degree their corporate behavior was deemed responsible by unbiased third parties performing in-depth, quantitative analysis of different kinds of corporate issues.

At the same time as Jantzi was developing his model in 1992, Fama and French published their study called ”The Cross-Section of Expected Stock Returns” (E. F. Fama & French, 1992), where they concluded that cross-sectional data on the risks identified as material to returns allowed investors to predict the return distributions of assets. When Fama and French published their asset pricing model, similar cross-sectional data was just developing in the ESG sphere, and foundations being laid for ESG integration being used by investors to recognize possibly more material cross-sectional for predicting returns of assets. (Sherwood & Pollard, 2018, 12-13)

OECD estimated that in 2020, the level of SRI investing in the US was over 20 % of professionally managed assets or about 11 trillion USD and 17 trillion USD in Europe (Boffo & Patalano, 2020, 15). Despite all the research and investor interest regarding SRI, the concept itself currently lacks an agreed-upon definition (Busch, Bauer, & Orlitzky, 2016; Fowler & Hope, 2007; Sparkes, 2002). Nevertheless this, for example, the UN's Principles of Responsible Investing. A set of guidelines published by the UN in 2006. (PRI) defines Responsible Investing (RI) as considering environmental, social, and governance (ESG) issues when making investment decisions and influencing companies or assets (UNPRI, n.d.).

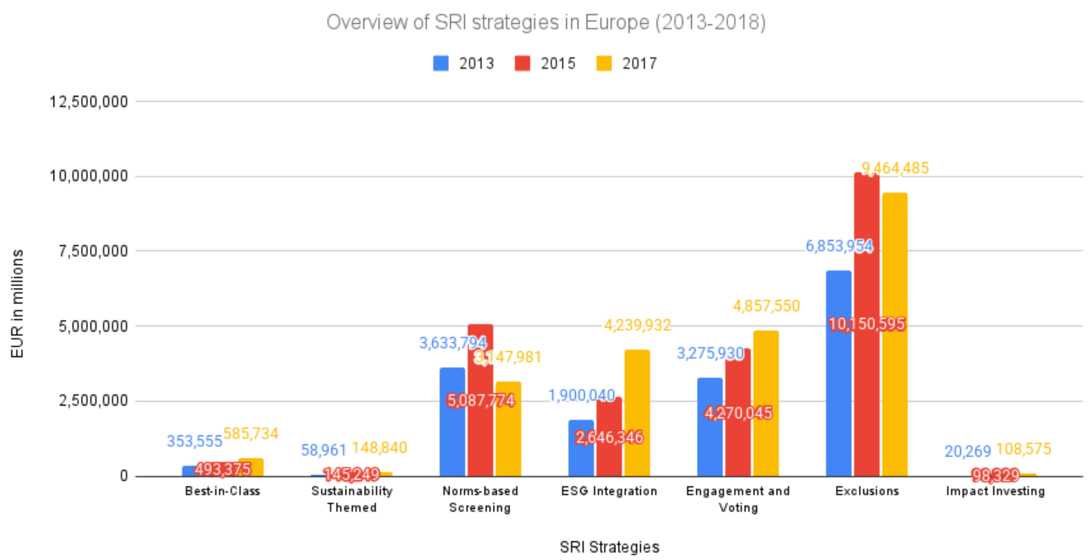

Figure 1. Overview of SRI Strategies in Europe (2013-2018) (Eurosif 2016; 2018).

2.2 Responsible Investment Strategies

Responsible Investing (RI)[1] is one of the many different concepts used to describe investment strategies, which attempt to meet certain ethical conditions while aiming to maximize financial returns. What the ethical conditions are and how the investment strategy attempts to meet them is up to each individual investor to decide.

[1]Other similar terms include, for example, Socially Responsible Investing (SRI), Ethical Investing (EI), and ESG investing (Bauer, Koedijk, & Otten, 2002)

There is no default or commonly accepted method for implementing a responsible investing strategy. Despite this, for example, Blowfield and Murray (2008) are able to recognize three different general investment approaches used in responsible investing: shareholder advocacy, community investment, and screening. The following gives a short description of the first two strategies. Screening will be covered in more detail in Section 2.3.1.

2.2.1 Shareholder advocacy

Shareholder advocacy is a strategy in which investors acting as shareholders attempt to positively influence corporate behavior. According to Schueth (2003), this may mean, for example, engaging in a dialogue with companies about issues of concern, as well as submitting and voting proxy resolutions. Divestments can also be seen as a form of shareholder advocacy.

2.2.2 Community Investment

Another form of responsible investing is community investing, where an investor may typically invest directly into a community institution which they consider to be providing a valuable service to the community. Frequently, community investments hope to support local economic development, often in poorer or otherwise disadvantaged communities. (Schueth, 2003)

2.3 ESG Investment Strategies

2.3.1 Screening

Screening can be either negative / exclusionary screening or positive / best-in-class screening. Negative screening may involve, for example, excluding companies belonging to specific industries or activities which are deemed unacceptable.

According to Caplan, Griswold, and Jarvis (2013) ESG investments began with negative screening, and according to Sherwood and Pollard (2018, 58) negative screening or divestments remain at the forefront of ESG investment implementation methodologies. In negative screening, according to Sherwood and Pollard (2018, 58), exclusion requires company-specific analysis for the investor to assess whether a particular company's operations are in accordance with the investor's beliefs, policies, or inclusion criteria. According to a 2012 report by the Global Sustainable Investment Alliance (2013), negative screening was the most widely used strategy in terms of managed assets.

In positive screening companies or countries with good or improving ESG performance relative to sector peers are over-weighed. According to Sherwood and Pollard (2018, 92), this can also be a method for systematic integration of ESG. Investors may choose, for example, to systematically integrate a universe of securities into their portfolio by using positive ESG ratings as a filtering factor. Positive screening can also be referred to as ESG tilting, as an investor is essentially putting more weight on their portfolio according to the quality of ESG ratings (Sherwood & Pollard, 2018, 92).

2.3.2 ESG integration

ESG integration systematically embeds environmental, social and governance considerations – together with other relevant non-financial metrics – into investment analysis and decision making (Sherwood & Pollard, 2018, 91). As Boffo and Patalano (2020, 8) note, it goes well beyond simple positive screening: when fully embraced, ESG integration permeates the entire investment culture, becoming an indispensable component at every stage of the investment process. Some organizations like Eurosif 2021a have developed methodologies to further classify sustainability-related investments.

2.3.3 Impact investing

In impact investing, investors can include factors such as how investing in a company with sustainable business policies may benefit the environment, in addition to considering how the investment may provide long-term advantages to the return potential of the investment (Sherwood & Pollard, 2018, 100). According to Sherwood and Pollard (2018, 100), themes can also be thematic in nature and aligned with, for example, general impacts or isolated to a particular issue or cause. These themes may be global or country-specific, for instance, or related to an impact on a certain culture, socioeconomic status, or population group. ESG investment with a thematic focus can also be seen as its own ESG investment strategy.

Table 1. ESG Incorporation strategies (adapted from UN PRI 2023; 2025).

Strategy	Definition
ESG Incorporation	Umbrella for three complementary approaches—integration, screening and thematic investing—typically reinforced by active ownership.
ESG Integration	Systematic use of financially material ESG data in research, valuation and portfolio construction to sharpen risk–return outcomes.
Screening	Rule-based ESG filters that determine whether securities, issuers or sectors are investable; can be positive, best-in-class, norms-based or negative.
Positive & best-in-class screening	*Positive*: admit assets that meet stated ESG thresholds. *Best-in-class*: admit leaders that rank above peers on those metrics.
Negative screening	Avoid or exclude securities that breach chosen ESG standards – e.g. controversial industry, ESG controversies, carbon intensity.
Norms-based screening	Require compliance with global norms (e.g., UN Global Compact, OECD Guidelines); exclude or engage on violations.
Thematic investing	Allocate investments to long-term ESG megatrends (renewables, inclusive growth etc.) via dedicated funds, indices.
Stewardship	Using investor rights to protect and enhance long-term value for beneficiaries and society.
Impact investing	Invest with the explicit intention of generating positive, measurable social and/or environmental outcomes alongside financial returns.

2.3.4 ESG Momentum

ESG momentum strategy is a combination of a more traditional momentum strategy presented by Jegadeesh and Titman (1993) and integrating into it a selection criteria for companies with improving ESG score development. The term momentum refers to an anomaly recognized by Jegadeesh and Titman (1993), where companies that have done well continue to do well, while companies with poor past performance will continue to do poorly. Jegadeesh and Titman (1993) found in their study that, based on past performance, shorting companies with poor performance and taking long positions in well-performing companies led to abnormal returns. Jegadeesh and Titman found that this

		Basic ESG	Advanced ESG	Impact-Aligned	Impact-Generating
Investment objective		Integration of ESG	Systematic analysis & incorporation of ESG factors	Align with positive impacts on environment and/or society	Measurable contribution to positive real-world impacts
Investment process	Investment Approach	Binding negative or positive screening	Binding negative & positive screening (≤80% of initial universe investable)	Binding negative & positive screening for assets with positive impact	Exclude non-transformable activities & use stewardship or provide new capital to assets to generate measurable positive impact
	Performance measurement	-	Measurement of ESG performance	Measurement of company impact	Measurement of company impact & investor contribution
Ambition Level		Low	Moderate	Medium	High
Investment focus					Double materiality

Figure 2. Eurosif's methodology to classify sustainability-related investments (2021b).

strategy produces significant positive returns over 3- to 12-month holding periods. The momentum anomaly recognized by Jegadeesh and Titman is covered in more detail in chapter 3.2.

Similarly, in ESG momentum strategies, Jegadeesh and Titman's approach is adapted by also considering the past ESG score performance of companies. In particular, attention is given to whether a particular company's ESG scores have been improving or declining. This also sets ESG momentum strategies apart from more traditional SRI strategies, where absolute ESG scores are typically used. In contrast, ESG momentum strategies aim to find clues about companies' future financial performance by looking at ESG score improvement or decline. Consequently, ESG momentum strategies typically invest not in the companies with the best ESG scores, but in those companies that have improved their ESG scores the most.

The concept of ESG momentum is a more recent innovation and seems promising. For example, the work of Giese, Lee, Melas, Nagy, and Nishikawa (2019) seems to suggest that at least companies with improved MSCI ESG ratings outperformed companies with declining ratings since the 2008 financial crisis. Similar results were found by Khan, Serafeim, and Yoon (2016), who also used MSCI ESG ratings to generate a dataset of customized ESG scores and performed regression analysis of stock returns with changes in the ESG score. Their findings suggested that there is significant predictive power behind ESG momentum for stock returns.

Nagy et al. in their 2016 paper compares the returns of two different ESG strategies, both of which use MSCI ESG ratings data. The first strategy is a more traditional ESG tilt strategy, which overweights stocks with higher ESG ratings, and the second is an ESG momentum strategy, which overweights stocks that have improved their ESG ratings in recent periods.

This study hopes to replicate the results found, for example, by Giese et al. (2019) and Khan et al. (2016) by examining the usefulness of the momentum of the ESG rating in predicting stock returns while controlling for other factors, such as size and book-to-market ratio, which are traditionally included in a Fama-French analysis. This will be covered in more detail in Section 5. Previous studies on ESG momentum and their results will be covered in more depth in section 4.

Despite their criticisms, Bruno et al. (2022) provides a comprehensive overview of the leading methods to create ESG and momentum-driven strategies. This overview is detailed in Table 3 below. For a clearer version of this table, see the Appendix 6.4.

Type of ESG Strategy	Papers	Results	Stock Universe
ESG Overall and Component Alpha	Deconstructing ESG Ratings Performance (Giese, Nagy and Lee, 2020)	Long/short portfolios from sorting on ESG, E, S, G (and other more specific key issue scores) come with **positive active return** and positive alpha.	MSCI AC World Index IMI (2013-2019).
	Foundations of ESG Investing: How ESG Affects Equity Valuation, Risk, and Performance (Giese, Lee, Melas, Nagy and Nishikawa, 2019)	"the **performance advantage** of higher ESG-rated companies is visible across the entire universe"	Several MSCI universes including Europe and US
ESG Momentum Alpha	Can ESG add alpha? An analysis of ESG Tilt and Momentum Strategies (Nagy, Kassam and Lee 2016)	ESG momentum strategies generate **positive alpha.**	MSCI World Index (2008-2015)
	How Markets Price ESG: Have Changes in ESG Scores Affected Stock Prices? (Giese and Nagy, 2018)	A long/short ESG momentum strategy shows a **positive alpha.**	MSCI World Index (2009-2018), MSCI Emerging market index (2013-2018).
ESG Combined Alpha	ESG for All? The Impact of ESG Screening on Return, Risk and Diversification (Verheyden, Eccles and Feiner, 2016)	Excluding stocks with lowest ESG scores leads to **improved returns,** lower volatility, and lower tail risk.	Large and Mid Cap Global and Developed (2010-2015).

Figure 3. Overview of papers which find positive alpha from ESG strategies (Bruno et al., 2022, 1).

As noted above, Bruno et al. were doubtful about recent research that aimed to find a positive alpha in ESG strategies. They concluded that, after correcting for sector and factor exposures, downside risk, and attention shifts, these strategies did not demonstrate significant outperformance. For example, enforcing sector neutrality significantly reduced the alphas reported for these strategies, suggesting that sector biases might play a role in their apparent effectiveness.

2.3.5 ESG Stability

The latest addition to these strategies is a strategy presented by Magnani et al. (2024), which suggests that ESG stability. While ESG momentum focuses on changes in firms' sustainability profiles, ESG stability tries to capture the consistency of those profiles over time. The central idea is that firms whose ESG scores evolve smoothly – exhibiting few spikes from incidents, controversies, or policy reversals – may be rewarded by investors for their predictability. In contrast, companies with highly volatile ESG trajectories may reflect only ad-hoc initiatives, weak internal controls, or exposure to recurring controversies that can also translate into cash-flow uncertainty, reputational drawdowns, and a higher cost of capital.

Magnani et al. link this perspective to established asset-pricing intuitions. Firstly, it resembles a "defensive" or "quality" style of investing in that stable non-financial policies may reduce downside risk and left-tail exposure. Secondly, smoother ESG paths can also be interpreted as information-quality: fewer revisions and controversies imply lower uncertainty, which investors may price. Additionally, stability is complementary to price momentum and traditional factors. This is another style of ESG investing this study hopes to shed further light on and particularly to replicate the results of Magnani et al.'s results on ESG stability. More details about ESG stability as a factor and Magnani et al.'s study is given in section 4.6.

3 Theoretical framework

This section of the thesis aims to frame the research question in a theoretical framework. A central factor in this framework is the Efficient Market Hypothesis and questions relating to it, which will be covered in the first subsection. This chapter also introduces some of the most commonly used portfolio performance measures. A brief overview of how to implement a momentum trading strategy is also given in the last section of this chapter. This section begins with an introduction to the Efficient Market Hypothesis and Modern Portfolio Theory (3.1), which help explain what portfolio outperformance or underperformance means. Subsequent sections cover typical

3.1 Efficient Market Hypothesis

The Efficient Market Hypothesis (EMH) was originally presented by Eugene Fama in his 1970 paper and has since become a cornerstone of modern financial theory. The EMH suggests that stock market prices fully reflect all available information. EMH can be categorized into weak, semi-strong and strong forms, each differing in the extent to which they believe that markets efficiently incorporate information into prices (Bodie et al., 2019, 335).

This work was based on Fama's earlier paper published in 1965 on random-walk theory, which concluded that successive price changes are independent and identically distributed random variables, indicating that past prices contain no useful information for predicting future ones (E. F. Fama, 1965). However, EMH does not suggest that prices are always ”correct” or that traders never fail to understand the information available. According to Bodie et al. (2019, 335), the EMH only suggests that, despite possible mispricings, prices on average should reflect the available information given sufficient time.

In the weak form of EMH, it is assumed that all previous trading information is reflected in stock prices. This implies that historical data, such as past prices, trading volume, or short interest, are already priced in (Bodie et al., 2019, 335-336). Consequently, methods such as technical analysis, which attempts to predict future prices based on historical data, would be ineffective in such markets.

In the semi-strong form of EMH, all publicly available past information regarding a company prospects has already been included in the price. This information includes everything publicly available to investors. In addition to past trading information, this includes information about, for example, fundamental data on a company's product line, quality of management, patents held, earnings forecasts, and balance sheet composition. (Bodie et al., 2019, 336)

The strong form of EMH goes farthest in defining all available information. The strong form asserts that stock prices reflect all information, public and private, past and present. Consequently, the strong form of EMH suggests that even, for example, insider trading, should not consistently outperform the market. The strong form of EMH has been widely tested and critiqued in the financial literature, for example, by Jaffe (1974).

Since his 1970 paper, Fama has adjusted his original formulation of the Efficient Market Hypothesis, taking into account accumulated research testing the various forms of EMH. In his paper Fama (1991, 1576) changes the original weak-form tests, which were concerned with how well past returns predict future returns, to ”cover a more general area of tests for return predictability”, including forecasting returns with variables such as dividend yields and interest rates, as well as considering the cross-sectional predictability of returns, i.e., tests of asset pricing models and the anomalies uncovered by tests.

Fama (1991), in addition, renames the semi-strong form tests of EMH as event studies. These have become a widely adopted methodology to measure the economic impact of a wide range of events. Strong-form tests are renamed as tests for private information, as, for example, Jaffe (1974)'s results have shown that private information can lead to abnormal returns and research by Sheyhun (1986) that corporate outsiders cannot profit from this insider information.

However, more recent studies challenge the absolute applicability of EMH in the context of non-financial data such as Environmental, Social, and Governance (ESG) information. Giese et al. (2019) find that while traditional financial data may be incorporated efficiently, factors related to ESG often show a delayed impact on stock prices. They suggest that ESG improvements (or ”momentum”) can yield positive returns as markets gradually integrate this information. This perspective offers a nuanced view of the semi-strong form of EMH, indicating that non-financial information may be assimilated over time rather than instantaneously.

3.1.1 Behavioral Finance and ESG Momentum

Another important perspective comes from behavioral finance, which challenges EMH by acknowledging that investors may not always act rationally. Shefrin (2002) provide insights into how cognitive biases, such as overconfidence and confirmation bias, influence investment decisions.

Li, Watts, and Zhu (2024) examine how retail investors react to news events related to ESG and find that their trading is primarily motivated by the financial implications of these events rather than ethical considerations. However, the study highlights that retail investors’ responses are more pronounced for high-attention events, such as those with significant media coverage. This suggests an attention-based behavioral component, where investors may prioritize ESG information when it is highly visible, aligning with behavioral finance theories that emphasize attention constraints in decision-making. This selective response could contribute to ESG momentum as investors gradually incorporate financially relevant ESG improvements.

These perspectives collectively form a theoretical framework that supports the potential value in ESG momentum. Although EMH implies that markets are generally efficient, the gradual integration of ESG information and investor biases challenges this view. Thus, both EMH and behavioral finance offer valuable lenses through which to understand the potential of ESG momentum strategies in modern investing.

3.1.2 Tests of the Efficient Market Hypothesis

Early tests of the EMH focused on the weak form and, in particular, on testing the efficacy of technical analysis. Essentially, these tests revolved around the question of whether it would be possible for investors to find useful trends in past prices of assets, which could be used to gain abnormal returns. (Bodie et al., 2019, 347)

For example, Conrad and Kaul (1988) and Lo and MacKinlay (1988) attempted to test the weak form of EMH over short time horizons by examining the serial correlation of weekly returns of NYSE stocks to find whether a positive serial correlation could be detected. Measurement of serial correlation of stock market returns in stock prices is a way to discern trends in stock prices and allows to analyze the tendency for stock returns to be related to past returns (Bodie et al., 2019, 347). Conrad and Kaul, as well as Lo and MacKinlay found evidence of momentum or positive serial correlation over short horizons, meaning that positive returns tended to follow positive returns.

As mentioned in section 2.3.4, Jegadeesh and Titman (1993) found evidence of momentum over longer horizons by using 3- and 12-month holding periods. Jegadeesh and Titman concluded that they could detect a momentum effect, where good and bad performance of stocks continued over time. In particular, Jegadessh and Titman concluded that while individual stocks predictions remained highly unreliable using their methods, constructing portfolios of recently best performing stocks seemed to outperform the market consistently enough to provide profit opportunities to investors.

Although EMH tests over short to intermediate time horizons have found evidence of positive serial correlation (momentum) in stock market prices, EMH tests over longer time periods, for example, Fama and French (1988), have found the opposite; evidence of negative long-term serial correlation in the performance of the aggregate market. According to Bodie et al. (2019, 347), these results have resulted in what he calls the ”fad hypothesis”. This hypothesis states that stock markets may overreact to news, for example, and these overreactions lead to positive serial correlation over short time horizons and to poor performance over longer horizons, when prior overreaction is corrected, leading to negative serial correlation over long horizons.

Although alternative interpretations of these results have also been given in the research literature, many studies have found that over long horizons extreme performance in securities tends to reverse itself. For example, De Bondt and Thaler 1985 and Chopra, Lakonishok, and Ritter 1992 were able to distinguish a strong tendency in both well-performing and poorly performing stocks to undergo a significant reversal in subsequent periods. This effect has been termed the reversal effect, where previously poorly performing stocks tend to improve their performance, and well performing stocks tend to do more poorly. One way to interpret this behavior is that, for example, short-term overreactions to news may lead to long-term reversal as corrections occur.

These and similar findings in the research literature have been hard to reconcile with the Efficient Market Hypothesis and are called anomalies in the literature. Anomalies are predictable and seemingly continuing trends that should not occur in an efficient market as predicted by EMH (Bodie et al., 2019, 349-353). thesis. According to Schwert (2002), these anomalies may indicate either market inefficiencies or inadequacies in the underlying asset-pricing model being applied. Some common types of anomalies will be covered in the following section 3.2 with an emphasis on momentum as that is the main interest of this

3.2 Market anomalies

At the end of the previous section 3.1.2, we presented a brief definition of market anomalies according to Bodie 2019. These anomalies often function as tests for the semi-strong form of the Efficient Market Hypothesis. Before delving further into the various anomalies, this section offers a brief overview of the challenges associated with interpreting the results of these tests.

A main difficulty in interpreting these tests highlighted by, for example, Bodie et al. (2019, 349), is that these risk-adjusted returns tests are, by definition, joint tests of the efficient market hypothesis and the risk adjustment procedure employed. This means that if, for example, a given trading strategy based on an anomaly seems to provide abnormal returns, we must weigh whether the risk-adjustment procedure has worked or rejecting EMH. Furthermore, for example, Schwert (2002) points out that after an anomaly has been documented and analyzed in the research literature, it often appears to disappear, reverse, or attenuate.

3.2.1 The Small firm Effect

The phenomenon known as the size effect or the small firm effect was identified by Banz (1981) and Reinganum (1981). They observed that companies with lower capitalizations on the NYSE seemed to yield higher average returns than the Capital Pricing Model anticipated. However, since its identification, the small-firm anomaly largely seems to have faded.

As EMH posits that asset prices fully reflect all available information, it follows that achieving higher returns than the market through stock selection or market timing should not be possible, at least not consistently. However, ESG momentum strategies attempt to capitalize on the persistence of stock performance based on ESG selection criteria. These strategies challenge the notion of market efficiency by suggesting that markets can systematically undervalue or overvalue stocks based on their ESG performance, leading to potential alpha generation. Or, more simply, as put by Bodie et al. (2019, 341), deviations from market efficiency (EMH) may offer profit opportunities for the well-informed trader.

3.2.2 Momentum anomaly and ESG

The momentum anomaly, discovered by Jegadeesh and Titman in their seminal 1993 paper, refers to the observation that stocks that have performed well in the past tend to continue outperforming in the short-term future, and vice versa for poorly performing stocks. This phenomenon challenges the efficient market hypothesis, suggesting that stock prices do not fully reflect all available information immediately. Jegadeesh and Titman demonstrated that a strategy of buying past winners and selling past losers could generate significant abnormal returns. This discovery has had profound implications for finance, particularly in the development of quantitative trading strategies.

For example, Nagy et al. (2016) found that a strategy based on ESG tilt and momentum outperformed market benchmarks. This finding directly contradicts EMH by demonstrating that a systematic strategy based on publicly available ESG information may be able to generate excess returns. Nagy et al. findings may suggest that the markets may be slow in assimilating ESG-related information into stock prices.

Nagy et al. 2016, 5 distinguish an ESG momentum strategy from a strategy that merely favors companies with high past ESG performance (also known as an ”ESG tilt” strategy). According to Nagy et al., ESG momentum is more short-term in nature and aims to link stock performance to changes in a company’s ESG scores over time. Unlike ESG tilt, this approach does not seek to enhance the overall ESG profile of the portfolio. Instead, it focuses on creating portfolios with companies that show the greatest improvements in ESG scores.

3.3 Portfolio evaluation metrics

This section covers some of the asset pricing models commonly used in the research literature. These include Capital Asset Pricing Model (CAPM) described by Sharpe (1964) and further developed by Lintner (1965) and Mossin (1966), the three-factor model by E. F. Fama and French (1992). Extensions of the Fama and French three-factor models, such as the Carhart (1997) four-factor model, which took into account momentum, are also included. Later revisions by Fama and French, such as the five- (2015) and six-factor models (2018) are also included, as these will be the base of the research methodology applied in this thesis.

3.3.1 CAPM

Traditional finance theory rests on the assumption that investors are rational and behave homogenously in attempting to maximize the returns of their investments. Investors do this by constructing their portfolios in a way that yields the highest possible return with their level of risk tolerance.

The CAPM serves as the foundational framework for all subsequent factor models discussed in this thesis. Because it underpins many of the fundamental principles in later theories, CAPM is arguably the most theoretically significant. Rooted in Modern Portfolio Theory (MPT), which Markowitz 1952 introduced, the CAPM focuses on the risk-return trade-off that investors must consider when evaluating investments. Markowitz’s mean-variance theory provided investors with a way to build optimal portfolios through diversification, aiming to maximize expected returns for a given level of risk tolerance.

CAPM builds upon Markowitz's work by explicitly addressing the relationship between the return on an individual asset and the performance of the broader market. To do this, CAPM introduces the concept of beta (β), a measure of an asset's sensitivity to market movements, thus defining a linear relationship between expected returns and systematic risk. According to CAPM, the expected return for an asset is a function of its risk relative to the market, with beta representing that risk. The model thus offers a method to calculate the expected returns, suggesting that investors should be rewarded based on the level of risk they assume in their investments (Lintner, 1965; Sharpe, 1964). Additionally, this implies that it has implications for asset prices themselves. An asset with a riskier payoff should have a lower price than an asset with less risky payoffs because a lower price is necessary to yield the higher returns demanded by investors as a risk premium. In equilibrium, then, asset prices should clear the market. (Grobys, 2014, 1)

The CAPM provides a structured approach to understanding how expected returns correlate with risk, reinforcing the principle that investors should receive compensation for the additional risk they take on. This concept has far-reaching implications, serving as the backbone for more complex models and playing a critical role in the theory and practice of portfolio management. The general nature in which CAPM helps to understand the relationship between risk and expected return has led to its widespread application in portfolio construction, performance measurement, risk assessment, and capital budgeting.

CAPM, like MPT, assumes that investors are rational, demand a higher return for taking on more risk, and that investors can minimize the unsystematic risk of their investments through diversification. More formally, the CAPM states that all risky assets should reward the investor with an excess return, which can be defined as $E(R_i) - R_f > 0$, i.e., the expected return on a risky investment in asset i should be positive compared to an investment on a risk-free asset R_f.

The CAPM equation is as follows:

$$E(R_i) = R_f + \beta_i \times (E(R_m) - R_f) \tag{3}$$

- $E(R_i)$is the expected return on an asset i
- $R_f$is the risk-free rate, often derived from government bold yields
- β_i is the beta of asset i, indicating its exposure to systematic risk
- $E_(R_m)$ is the expected return of the market
- $E(R_m - R_f)$ is the Market Risk Premium (MKT), representing the additional return for taking on market risk

Equation 3 illustrates a positive correlation between the expected return of the asset i and its beta, β_i. Here, β_i represents the responsiveness of the asset i to market movements, defined as the gradient of the regression line between R_i and R_m (Black, 1972). Furthermore, this equation indicates that an increase in market risk corresponds to a higher expected return, or more specifically, $E(R_i) - R_f$. The specific risk associated with the asset i, denoted as β_i, is calculated by the ratio of the covariance of the returns on the asset i and the market portfolio (Cov(R_i, R_m)) to the variance of the market portfolio (Var(R_m)), as shown in equation 4. Typically, β_i is determined through a linear regression analysis of (3) using historical return data.

$$\beta_i = \frac{\text{Cov}(R_i, R_m)}{\text{Var}(R_m)} \tag{4}$$

Figure 4 describes investment opportunities according to the Capital Asset Pricing Model. In the figure, the horizontal axis represents portfolio risk, which is a measure of the standard deviation of portfolio return, while the vertical axis displays expected return. In the image, the curve abc is the minimum-variance frontier, which includes various combinations of expected returns of risk assets and risk that minimize the variance in returns at different levels. These portfolios do not include risk-free borrowing or lending. Without risk-free borrowing and lending, only portfolios above point b and along abc are mean-variance-efficient, as these will maximize expected return given the return variances for these portfolios. The figure also exemplifies the trade-off between risk and expected return for minimum-variance portfolios. Investors hoping to get a higher expected return, for example, at point a, must also accept the higher volatility associated with such a portfolio. At point T, the investor can receive intermediate expected return with lower volatility. (E. F. Fama & French, 2006, 26-27)

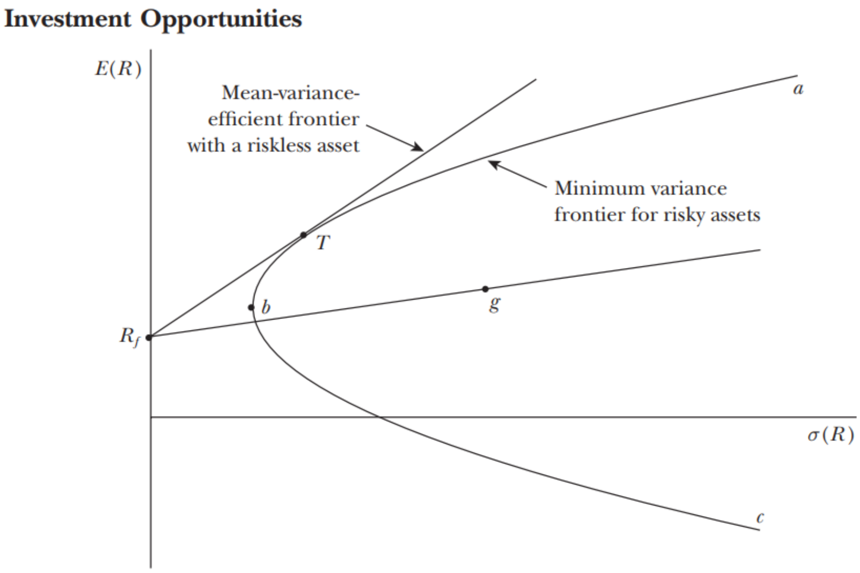

Figure 4. CAPM investment opportunities (E. F. Fama & French, 2006, 27).

The CAPM continues to be widely used in finance due to its simplicity and based on the assumption that the expected rate of return for a security can be derived from the risk-free rate, the expected market return, and the market beta. Bodie et al. (2019, 278) present the following simplifying assumptions that the CAPM model relies on. These assumptions are based on individuals acting identically with the notable exception of their initial level of wealth and level of risk aversion.

1. Individual behavior
   1.1. Investors are rational, mean variance optimizers.
   1.2. Their common planning horizon is a single period.

1.3. Investors use identical input lists, that is, homogeneous expectations. This is consistent with the fact that all relevant information is publicly available.

2. Market structure
   2.1. All assets are publicly held and trade on public exchanges.
   2.2. Investors can borrow or lend at a common risk-free rate and can take short positions on traded securities.
   2.3. No taxes.
   2.4. No transaction costs.

Although CAPM has been extremely influential and a building block for many other theories, it has also faced criticism. Some of these include that beta may not accurately capture an asset's true risk, its assumptions of market efficiency, and a linear risk-return relationship (Amihud & Mendelson, 1986; Blume, 1975; Campbell & Cochrane, 1999; Campbell & Vuolteenaho, 2004; E. F. Fama & MacBeth, 1973; Jensen, Black, & Scholes, 1972). These critiques have led to the development of more complex models, such as the Fama-French factor models, which address some of CAPM's limitations.

Sharpe ratio is a performance measure often associated with CAPM. Introduced by Sharpe (1966) as a reward-to-variability ratio, which compares the excess return of an asset to the variability (standard deviation) of said excess returns. As such, it enables comparison between portfolio returns, strategies, or even asset classes by simply looking at the returns and the amount of risk associated with these returns. Thus, a higher Sharpe ratio signals more reward per unit of risk. The equation for Sharpe ratio can be defined as follows:

$$\text{Sharpe ratio} = \frac{R_i - R_f}{\sigma_i} \tag{5}$$

- $R_i =$ Return on portfolio or asset i
- $R_f =$ Risk-free rate of return
- $\sigma_i =$ Standard deviation of the portfolio's periodic returns (i.e., typical up-and-down swing)

Another commonly used risk-adjusted performance measure is Jensen's alpha is also an extension of CAPM. The metric introduced by Michael Jensen (1967) is a metric which measures whether an investment earned more (or less) than the return justified by its exposure to market risk, according to CAPM. When a portfolio has positive alpha, it has outperformed the market index or benchmark. In summary, Jensen's alpha gives the average return of the portfolio over and above that predicted by CAPM, while considering the portfolio's beta and the average market return (Bodie et al., 2019).

$$\alpha_p = \overline{R}_p - [\overline{R}_f + \beta_p(\overline{R}_M - \overline{R}_f)] \tag{6}$$

When used in a multifactor market context, for example, with the Fama-French 3-factor model, an alternative formulation can be given, where s_p is the loading on the SMB and h_p the loading on the HML portfolio:

$$\alpha_p = \overline{R_p - R_f} - \beta_p \left(\overline{R_M - R_f}\right) - S_p\overline{R}_{\text{SMB}} - h_p\overline{R}_{HML}, \tag{7}$$

3.3.2 Fama-French three factor model

The Fama-French three-factor model (1992) was derived from a necessity to extend the CAPM and produce results that better reflect reality. For example, Jensen et al. (1972) and E. F. Fama and MacBeth (1973) found a positive linear relationship between average stock returns and β in the pre-1969 period. However, later researchers find that the relation between β and average returns seems to disappear in more recent periods, for example, during 1963-1990 (see e.g., E. F. Fama & French, 1992; Lakonishok & Shapiro, 1986; Reinganum, 1981). As a result, Jensen et al. suggest that potentially only the β_i factor alone was unable to capture all systematic risk and the CAPM model should be expanded and at least one additional factor may need to be included in the CAPM model.

The three-factor model by Fama and French (1992) extends the CAPM by tackling some of its shortcomings, particularly with respect to elucidating asset returns. While the CAPM was limited to market risk, Fama and French expanded their model to include additional factors that more effectively account for fluctuations in stock returns. These factors are the Size Premium (SMB - Small Minus Big) and the Value Premium (HML - High Minus Low). These factors can also be interpreted as risk premiums; the SML is the risk premium associated with size and the HML is the risk premium associated with the value versus growth characteristics of a company.

The size premium attempts to capture the out-performance of smaller companies over larger ones. This was based on the observation that smaller firms tend to outperform larger firms on a risk-adjusted basis. The size premium is calculated as the difference in returns between a portfolio of small-cap stocks and a portfolio of large-cap stocks. The size effect was first documented by Banz 1981. The size premium measures the firm size through market capitalization, which is the market value of a company's outstanding equity.

The value premium, on the other hand, reflects the observation that value stocks tend to outperform growth stocks, i.e., based on the tendency of companies with high book-to-market ratios (value stocks) to outperform those with relatively low book-to-market ratios (growth stocks). The value premium is calculated as the difference in returns between a portfolio of high book-to-market stocks and a portfolio of low book-to-market stocks. The value effect was first documented by Basu in his 1977 paper. The value premium is measured using the book-to-market ratio, which is calculated by dividing the book value per share by the stock price.

The resulting model suggested by Fama and French is presented in Equation 8:

$$E(R_i) = R_f + \beta_{i,M} \times (R_M - R_f) + \beta_{i,\text{SMB}} \times \text{SMB} + \beta_{i,\text{HML}} \times \text{HML} + \epsilon_i \tag{8}$$

where

- $E(R_i$ is the expected return on portfolio i,

- R_f is the risk-free rate,
- R_M is the expected return of the market,
- $\beta_{i,M}$ is the sensitivity of asset i to market risk,
- $\beta_{i,\text{SMB}}$ is the sensitivity of portfolio i to the size factor,
- $\beta i,$ HML is the sensitivity of portfolio i to the value factor,
- ϵ_i is the

By including these factors, Fama and French attempted to enhance the CAPM by providing a more nuanced understanding of asset returns by considering also market, size, and value factors. Furthermore, the three-factor model recognizes that risks beyond general market movements can also impact investment performance, thus offering a broader framework for asset pricing and portfolio evaluation. In general, the Fama and French three-factor model can be seen as a methodology to recognize the most likely sources of systematic risk (Bodie et al., 2019, 321). Usually, the chosen variables are those that, in the past evidence, seem to have predicted average returns well. Later, this model expanded further into five- and six-factor models, which continued to refine the model and expand upon the insights from the original CAPM model.

Although the Fama-French three-factor model and its variations have become prevalent in empirical studies of security returns (E. F. Fama & French, 1996), these methods have faced criticism. A significant critique is that not all factors used in these models can be definitively linked to an investor's risk exposure. For instance, Black (1993) suggests that employing a Fama-French type method might result in what is known as data-snooping. This practice involves researchers combing through databases of security returns to find explanatory factors, potentially discovering patterns that are merely coincidental.

3.3.3 Carhart four-factor model

Motivated by the Jegadeesh and Titman (1993) findings and the three-factor model developed by E. F. Fama and French (1992), Carhart (1997) developed a variation of the Fama-French three-factor model, where he included momentum as a factor. Carhart's goal was to use his model as a tool to evaluate mutual fund performance.

The factor included by Carhart is called Winners Minus Losers (WML) and its construction is otherwise similar to the Fama and French factors.

$$E(R_i) = R_f + \beta_{i,M} \times (R_M - R_f) + \beta_{i,\text{SMB}} \times \text{SMB} + \beta_{i,\text{HML}} \times \text{HML} + \beta_{i,\text{WML}} + times\text{WML} + \epsilon_i \tag{9}$$

where

- $E(R_i$ is the expected return on portfolio i,
- R_f is the risk-free rate,
- R_M is the expected return of the market,
- $\beta_{i,M}$ is the sensitivity of asset i to market risk,
- $\beta_{i,\text{SMB}}$ is the sensitivity of portfolio i to the size factor,
- $\beta i,$ HML is the sensitivity of portfolio i to the value factor,
- ϵ_i is the

3.3.4 Five-factor and six-factor models

Fama and French (2015) decided to expand their original model with factors related to profitability and investment factors. Fama and French tried to better account for the fact that growth and value companies had seemed to have differing returns under certain economic conditions. Especially, the findings by Titman, Wei, and Xie (2004), which suggest that there is a negative relationship between firms increasing their capital investments and worse stock returns in the future. Furthermore, Novy-Marx (2013) had found that profitability, measured by gross profits to assets, had equivalent predictive power to the book-to-market ratio to predict in the cross section of average returns.

The new model incorporates a profitability factor known as RMW (Robust Minus Weak), which represents the differential returns of portfolios characterized by strong versus weak profitability. According to Fama and French, companies that exhibit higher operational profitability generally achieve higher average returns. In contrast, the investment factor, labeled CMA (Conservative Minimum Aggressive), demonstrates an inverse relationship. Companies that adopt a more conservative investment approach typically outperform those that engage in more aggressive investment strategies. The resulting model becomes the following equation 10:

$$\begin{aligned} R_i - R_f = \alpha_i + \beta_{i,M}(R_M - R_f) + \beta_{i,\text{SMB}}\text{SMB} \\ + \beta_{i,\text{HML}}\text{HML} + \beta_{i,\text{RMW}}\text{RMW} + \beta_{i,\text{CMA}}\text{CMA}\epsilon_i \end{aligned} \tag{10}$$

In this equation, unlike the earlier ones, the factor α_i has been included, but assuming that the five-factor model captures all the variation in expected returns, $\alpha = 0$. Alpha is the intercept for all securities (E. F. Fama & French, 2015).

According to Fama and French, the five-factor model was able to explain 94% the cross-sectional variance in observed returns. The latest revision to the Fama and French model is the six-factor model, which includes a momentum factor (UMD, Up Minus Down), often also shortened as MOM. This factor tries to account for the momentum anomaly found in Jegadeesh and Titman.

The momentum factor is constructed by ranking stocks according to their returns in the last 12 months, omitting the most recent month to circumvent the short-term reversal phenomenon. The MOM factor is established by calculating the return differential between the highest and lowest deciles. It represents the performance of a portfolio that takes long positions in stocks within the top decile of returns and short positions in stocks within the bottom decile. (E. F. Fama & French, 2018)

4 Previous studies on ESG momentum and stability

This section gives a short overview of the various approaches used to study ESG momentum in the research literature. As ESG momentum is a relatively new SRI strategy, there are not yet too many studies in which it has been tested. In addition to academic research, some white papers have been published on the ESG momentum strategy by MSCI and other organizations implementing this strategy, which mostly focus on testing the performance of this strategy but provide useful information on its applicability. This is by no means a comprehensive literature review of all papers focusing on the topic of ESG momentum, but mainly sets the background in the studies this paper is rooted on.

4.1 Nagy, Cogan, and Sinnreich (2013)

In 2013, Nagy, Cogan, and Sinnreich published the first academic paper on ESG momentum. Their goal was to study find out whether it would be possible to improve the ESG profile of a global equity portfolio without sacrificing performance. They use MSCI World constituents, MSCI's 10-point Intangible Value Assessment (IVA) ratings and the Barra GEM-3 risk model to back-test three different quarterly-rebalanced strategies between February 2007 to December 2012. These strategies include a (i) worst-in-class exclusion of all ESG laggards (according to the MSCI's IVA rating), (ii) a simple ESG tilt, which overweights high ESG stocks and underweights low ESG stocks, and (iii) an ESG momentum strategy, which overweighted companies, whose IVA scores had risen over the previous 12 months and underweighted those whose had fallen.

Throughout the time period, all three strategies succeeded in improving the average ESG ratings of the global equity portfolios generated, but the ESG momentum strategy generated the strongest risk-adjusted results. It achieved an annual risk-adjusted return of 0.35% with an information ratio of 0.97.

Through decomposition of their results, the authors found that the consistent underperformance of downgraded stocks rather than the outperformance of existing leaders drove most of the alpha. This, according to the authors, suggests that the markets are more reactive toward negative ESG news than toward positive rating upgrades.

The authors additionally looked at where their ESG-tilted portfolio added or lost value and found that the results were not uniform across industries and geographies. They found that especially in the EMEA region, where ESG regulation and disclosure standards are relatively strict, the risks of contamination are very high. Symmetrical selection effects appeared most strongly. In the EMEA, the effect is most strongly observed in the utilities and telecom sectors. In North America and the Asia-Pacific regions the results were messier and high-rated stocks did not always outperform nor low-rated underperform, and their tilt-strategy even hurt performance. The authors attributed this to the looser ESG rules and the thinner analyst coverage in these markets, which could mean that ESG information is not priced as consistently or as quickly in share prices.

4.2 Nagy, Kassam & Lee (2016)

Nagy continued the study of ESG tilt and momentum strategies in 2016 with Kassam and Lee. In this study, they focused on the earlier ESG tilt and ESG momentum strategy and allowed a tracking error of up to 2.5% to test whether either strategy could use ESG signals to deliver excess returns and alpha.

Using MSCI World constituents and an extended eight-year time window (2/2007–3/2015), both strategies beat the portfolio benchmark. The plain ESG tilt strategy, which used current ESG scores, had an annualized active return of 1.1%. However, the momentum strategy delivered 2.2% and performed more consistently. A significant part of this outperformance remained unexplained even after the authors removed well-known style factors, suggesting that the effect could be genuine. Through decomposition, the authors found that nearly 90% of the tilt strategy's active risk and two-thirds of momentum strategy's came from stock-specific sources. Both strategies also raised the ESG rating of the portfolios with the tilt strategy raising it by 4 points on MSCI's 0-10 ESG scale and momentum by an average of 1.7 points.

4.3 Verheyden, Eccles, Feiner (2016)

The Verheyden, Eccles and Feiner (2016) study focuses on the question of whether applying a simple ESG screen in a stock selection process leads to worse performance. To do this, they constructed two broad asset universes: Global All (large- and mid-caps from 46 developed and emerging countries) and Global DM (23 developed markets). To these universes, the authors apply three different ESG filters: (i) best-in-class removal of the bottom 10% and 25% of the bottom ESG scorers in each industry; (ii) exclusion of those firms, which breached the UN Global Compact Principles; (iii) readmission of previously excluded firms, whose ESG score has improved over the last 3-6 months (ESG-momentum).

Between 2010-2015, the authors found that the screens did not harm the risk and return metrics compared to benchmark. In the Global All universe, dropping the worst-ranked 10% or 25% of stocks raised the annualised returns by 0.30% and 0.21%; in Global DM, the 10% screen added 0.15% while the 25% version shaved 0.01%. All but the Global DM portfolio with a 25% filter had higher Sharpe ratios with the volatility and drawdowns remaining marginally lower than for their unscreened counterparts.

In conclusion, Verheyden, Eccles, and Feiner found that their ESG screens seemed to dampen the downside tails of the return distribution and that excluding the ”dirtiest” quartile removed many of the worst daily losers, shifting the return distribution toward positive territory and trimming extreme-loss skewness and kurtosis. The screens applied also did not raise the specific risks of the portfolios with the correlation to the parent index staying above 0.99. This suggests that applying a simple filter does not significantly harm diversification. The study shows that even without seeking alpha, like Nagy, Kassam, and Lee (2016), basic ESG ”hygiene” or filtering may not impose a performance tax and may in some cases even offer a small premium.

4.4 Giese, Lee, Melas & Nishikawa (2019)

Giese, Melas, and Nishikawa in their 2019 paper shift the focus from whether ESG performance can improve returns to how ESG performance improves returns. By drawing on theory of corporate finance they suggest three transmission channels - cash-flow, idiosyncratic risk, and valuation - though which ESG information could feed into share-prices. The authors test their theory by embedding MSCI ESG ratings into a discounted-cash-flow framework and then test each channel directly to seek for causal evidence.

Using over 1600 MSCI World constituents between 2007-2017, the authors found that across the decade the highest ESG-rated quintile generated higher profitability and dividend yields than the lowest quantile, supporting the cash-flow channel hypothesis. At the same time, strong ESG performers also exhibited smaller idiosyncratic risk tails, with fewer 95 % drawdowns, lower residual volatility, and lower return kurtosis, supporting the insurance-like (idiosyncratic) transmission channel. Additionally, the highest ESG-rated quantile had lower systematic volatility and beta, hence lower implied costs of capital, and traded with higher valuations (i.e., lower book-to-price and earnings-to-price ratios), which seemed to also confirm the valuation channel hypothesis.

Most significantly, the study tests the causality of whether upgrades in a firm's ESG rating presaged subsequent drops in systematic and stock-specific risk and led to higher valuations and whether downgrades had the opposite effect. These rating changes, which can also be termed ESG momentum, appeared to translate into return premia as well meaning that the top momentum quantile outperformed the bottom quantile over the 2007-2017 period, which corresponds with the earlier portfolio backtests carried out by Nagy et al. in 2016. The authors further showed that while these ESG signals were low intensity, they are long-lived relative to classic factors like price momentum, which could make them particularly suitable for strategic benchmarks and low-turnover mandates.

Especially compared with the aforementioned optimization papers by Nagy et al., which put a primary focus on whether portfolios tilted toward static or improving ESG scores could beat benchmark, the paper by Giese et al. contributed to the literature by digging deeper into the firm-level mechanics behind those portfolio results. Giese et al.'s evidence suggests that the alpha documented in the tilt and momentum studies is, in fact, rooted in a combination of superior cash generation, risk mitigation, and repricing of capital costs, rather than a standalone ESG factor as such. Therefore, the paper reinforces the case for integrating dynamic ESG metrics into both security valuation and index construction.

4.5 Chen & Yang (2020)

Other more recent notable studies on ESG momentum and related phenomena include Chen and Yang's (2020) paper, which tries to test ESG momentum as a behavioral signal rather than a risk factor in the Taiwanese stock market. The authors feel that this market is particularly suitable for this kind of testing, as it, unusually, lacks the price-momentum anomaly.

The authors used LSEG's (formerly Thomson Reuters) ESG scores and studied the period between 2010-2017. By forming 25 double-sorted portfolios monthly, the results indicated that buying past return winners with high ESG scores and shorting ”losers” with low ESG scores earns 1.3-2.3% a month for up to 18 months, after which returns seemed to reverse. This may be due to short-term exuberance or long-term corrections occurring according to the authors. In the study, the alpha survived CAPM and Fama-French controls and seemed to be mainly driven by the Environmental pillar. However, Chen and Yang's observation that stocks with positive ESG surprises seemed to outperform shortly after the news, only to see their excess returns dissipate over longer horizons, also highlights a possible reason, why detecting ESG alpha may be so difficult. Long holding periods may dilute or miss the transitory gains of positive ESG news events suggesting that for ESG alpha to be capturable, more frequent portfolio updates may be required. A strategy which rebalances annually based on ESG changes, for example, may miss bulk of the ESG momentum gains, which could have already occurred and faded, when a rebalance occurs.

A recent study by Kondouri, Landis & Pittis (2025) seems to support the findings of Chen and Yang. To address the possibility that annual portfolio rebalancing may be too slow to capture abnormal returns using ESG momentum, they use employ monthly rebalancing for their ESG momentum factor and industry analyses to backtest their ESG momentum strategies. With a mothly rebalancing cycle Kondouri et al., like Nagy et al. before them (2016), were able to capture a cumulative outperformance of about 2% per year. The long-short ESG momentum factor strategy by Kondouri et al. had similarly high risk-adjusted returns (monthly average 0.73% and Sharpe 0.7) and significant alpha. Kondouri et al. further confirmed their results by running a Newey-West test, which returned $t \approx 2.7$ for monthly returns, confirming the statistical significance of the findings. This result was also heteroskedastic and autocorrelation-consistent (HAC) with an HAC $t > 4$, which implies strong statistical significance even after correcting for heteroskedasticity and autocorrelation. The HAC value roughly translates to a two-sided p-value below 0.00005. (Magnani et al., 2024)

4.6 Magnani, Guidolin & Berk (2024)

Magnani et al. (2024) are also interested in ESG momentum, but shift the focus from returns to the cost of equity capital (in Europe). As mentioned earlier in 2.1 earlier studies by, for example, Hong and Kacperczyk (2009) have suggested that compliance with the CSR may impact stock returns through capital cost. In their cross-sectional tests Magnani et al. ask whether investors reward firms more for an upward slope (”strong”) or low volatility (”stable”) in ESG ratings. The study finds that while short-term momentum is, in fact, priced in, its effect on the ex-ante cost of capital varies sign; sometimes an improved ESG score is associated with a lower required return required and the firm’s funding becoming cheaper, but sometimes an improved ESG score can also be linked to a higher required return. Magnani et al. found that while ESG rating momentum delivered mixed pricing signals, ESG performance or rating stability, in contrast, reliably lowers a firm’s cost of equity. Especially, Magnani et al. (2024, 677) find that ESG momentum seems to be defined only in relatively short time windows of 1 and 3 months.

Magnani et al. developed a sample portfolio with long-short spread of ESG volatility, which bought low-volatility (stable) ESG scorers and sold high-volatility ones. This returned a robust alpha and unambiguously lowered the implied cost of capital. According to the authors, this result was maintained in the MSCI and Sustainalytics ESG rating data. These results suggest that a solid, controversy-free ESG profile with stable long-term oriented objectives may be more beneficial than one seeking ESG rating upgrades for short-term benefits. A stable ESG profile dependably lowers financing costs, and investors seem to reward companies for *predictability* in sustainability performance more than for bursts of recent improvement, which may prove short-lived.

Magnani et al.'s observation that investors reward companies for long-term sustainability performance appears to be in line with Koundouri and Landis (2023) study, where stocks that increased their ESG performance over a 2-year period tended to realize high abnormal results. Similarly, the study De Lucia, Pazienza, and Bartlett, which examined whether stronger ESG performance leads to better financial performance results, points to the fact that broad and strategic ESG commitments had clear and statistically significant links to higher ROE and/or ROA while short-term tactical measures aimed at improving a company's ESG score may reduce short-term profitability. This seems to track with Magnani et al.'s conclusion that investors reward companies for predictability in their ESG performance more than temporary and possibly short-lived improvements in performance.

Additionally, many studies that find the ESG momentum alpha often seemed to focus on specific regions, shorter horizons, or particular segments of the market, where the signal-to-noise ratio for ESG news is higher and legislation requires that sets clearer disclosure standards, for example. This means that some results may not be replicable in other segments or regions. For example, Alves, Krueger, and van Dijk examined the relationship between ESG changes and stock returns in a broad, global cross section, which covered over 16 000 stocks in 48 countries and did not find a robust relationship between ESG changes and stock returns. The contrast between the findings of broader studies, like Alves et al.'s, and studies focusing on a more specific region or strategy, suggests that methodology matters greatly and some of the ESG effects may be context-dependent or short-lived enough that at a global aggregate level, they get washed out by noise or arbitrage.

5 Data and methodology

This section covers the basic details of the data used in this project and the limitations imposed by it. In addition, an overview of the methodology applied in the thesis is provided. In Section 6 the results of the study will be given and explained. The attempt is to try and find out whether the momentum and / or stability of ESG scores could be a common risk factor in the equities market. Next, we will cover details about ESG data and their availability in this study, as well as other data used.

5.1 Data

5.1.1 ESG data

In this study, the constituents of the S&P 500 Index are used as the universe of stocks analyzed. The constituents of the S&P 500 are highly liquid and using it gives us compatibility with the Fama-French framework, where the FF5 extension was specifically developed for the U.S. equity market with factors estimated from U.S. stocks using U.S. accounting rules. An additional factor to consider is that using the methodology of Magnani et al. (2024) requires having access to monthly ESG data.

The data for this study came exclusively from LSEG's (formerly known as Refinitiv / Eikon / "Asset4 ESG") database. The main exclusion is that, for this thesis, I did not have access to LSEG's' monthly ESG data. Data collected from LSEG included equity prices, fundamental, sector/industry and information on, which stocks were part of the S&P 500 at different times. Due to the unavailability of monthly ESG data from LSEG, this data was collected from Yahoo Finance for the S&P 500 constituents during the period 2014-09-01 - 2024-07-01. Yahoo Finance in turn gets its ESG data from Sustainalytics, which is a recognized ESG ratings provider (Sustainalytics, 2020). Figure 5 illustrates changes in S&P 500 constituents between 2014-2024 using data from LSEG. As can be observed, there have been significant changes in S&P 500 constituents. To keep track of these, the monthly data on the prices, fundamentals, and ESG factors of these constituents were fetched each month in the period 2014-2024. Fundamental data, price data came from LSEG. Here month end adjusted close prices and market cap data was used to calculate returns and size-related factors for individual assets. This data was enriched by including sector information on each company (TRBIC).

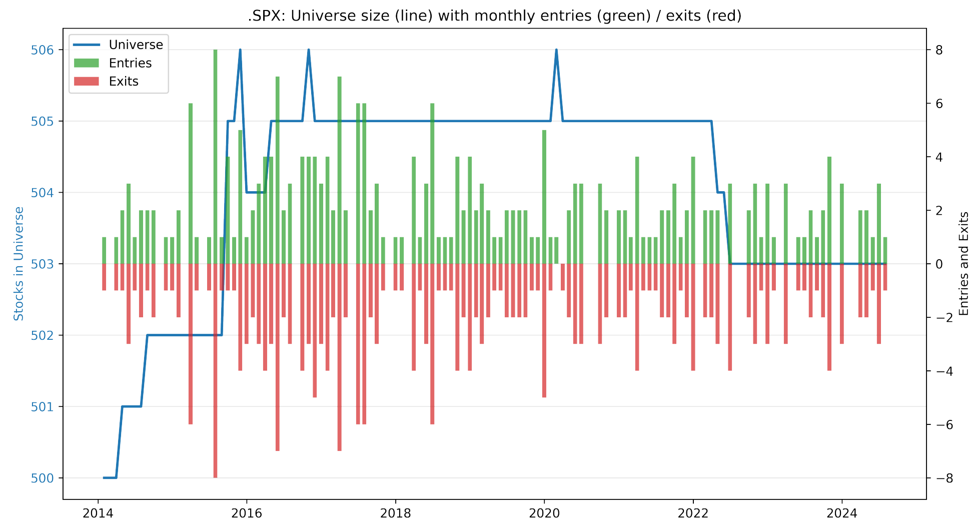

Figure 5. Entries and exits in S&P 500 (SPY point-in-time constituent data from LSEG).

ESG data itself can be problematic. As mentioned earlier in the Introduction, in recent years, there has been a proliferation of ESG ratings providers. However, the ratings methodologies and materiality assessments used by these ratings providers (see, e.g., LSEG, 2025; MSCI, 2025; Sustainalytics, 2024) can vary significantly.

A few studies have focused on studying the issues one needs to address or consider, when utilizing ESG data. Kotsantonis and Serafeim (2019, 51), for example, recognize four main limitations in the currently available ESG metrics:

1. Data inconsistency of raw data (metrics disclosed by firms)
2. Distortions are introduced by ”benchmarking” (or through differing definitions of peer groups)
3. ESG data imputation can be a problem (or not all models are created equal)
4. ESG data providers disagree a lot (and even more, surprisingly, when there is publicly available information)

The first problem concerns relates to how companies themselves disclose ESG-related data, which rating providers use to assign ESG scores. Frequently, companies disclose data using different metrics, different terminologies, and even different units of measure. To illustrate the problem, Kotsantonis and Serafeim selected a random sample of 50 Fortune 500 companies and found 20 different ways these companies reported their Employee Health and Safety data with several of the metrics measuring different things. Given that rating providers assign ESG scores by comparing a snapshot of performance of a given company in relation to a range of comparable values which define the best and worst possible performance for a given sample, this easily leads to a situation where relative performance is difficult if not impossible to compare in some areas. Table 2 illustrates the difficulty in comparing performance, when different metrics are used, for example, to measure employee health and safety.

Table 2. Various different units used to measure Employee Health and Safety in random sample of 50 large, publicly listed companies (Kotsantonis & Serafeim, 2019, 52).

Company	Metric	Unit
1	Number of accidents with fatal consequences	Number
2	Rate of injury per 200k hours worked	Number (ratio)
3	Occupational injury related fatalities	Number
3	Occupational injury related fatalities	Number
4	Lost-time incident frequency rate	Percentage
5	Injury rate	Percentage
6	Total case incident rate	Percentage

As Kotsantonis and Serafeim (2019, 52) define an ESG metric as a snapshot of performance that is assessed in relation to a range of values, which define the best and worst performance, which can be done either through peer group analysis or by assessing absolute levels performance based on some predefined ”optimal” level of performance on ESG metrics. This leads to the second primary source of discrepancies among ESG rating providers, which is the definition of the range of best and worst performance.

5.1.2 Defining a peer group

Rating providers are free to use discretion in how they define peer group, but typically either of two options are used. (1) Universal peer groups are a sample of companies across countries and sectors, for example, MSCI ACWI. When ESG performance is calculated with universal peer group as a benchmark, there will unavoidably be industry-level bias and possibly geographical differences due to differences in legislation, for example. The alternative option is (2) peer groups in industry, where a sample of companies in the same primary industry or sub-industry is used to benchmark ESG performance.

Regardless of which peer group definition is chosen, the range of performance observed in the peer group will become the yardstick against which individual company's performance is assessed. Therefore, differing definitions between ratings providers may crucially impact, how a company's ESG performance is ranked. This is particularly likely when different providers use different industrial classifications, such as the Global Industrial Classification System (GICS), MSCI IVA industries, or the Bloomberg Industrial Classification System (BICS), or when sample periods and, therefore, the peers are different.

Assessment of diversified businesses is similarly difficult. Many large companies may receive revenue streams from various business units, which may not fall under their primary industry classification. It should be noted that no agreed method as to how diversified businesses like these should be assessed in terms of ESG performance nor which ESG issues are material to them.

There have been many studies on the divergence between ESG scores assigned by different providers. Kontsantonis and Serafeim's (2019) study recognized three different sources of ESG rating divergence between providers: (i) Scope divergence", where the ratings are based on different sets of attributes; (ii) "Measurement divergence" occur when rating providers use different indicators to measure a given attribute (see, e.g., Table 2); (iii) "Weight divergence" are caused by rating providers having differing views on the "materiality" or importance of given attributes (e.g., provider A assigns indicator x higher weight than provider B). The authors found that in their sample measurement, divergence contributed 56%, scope divergence 38%, and weight divergence 6% to the observed divergence. Kotsantonis and Serafeim also found that measurement divergence is partly driven by the so-called ”rater effect”, where a firm that receives a high score in one category is more likely to receive high scores in other categories as well from the same rater. The rater effect was later also observed by Berg, Kölbel, and Rigobon in their 2022 paper, where they note that the rater effect may introduce correlated errors that bias event study results.

Liang and Renneboog summarize that the ratings differ due to different providers applying different set of indicators, weights, and qualitative adjustments. These idiosyncracies will also develop over time as, for example, the provider re-evaluates the importance or materiality of an ESG issue for a company or a given sub-industry. Therefore, rating changes may even reflect a rating provider’s own materiality thinking evolving rather than any corporate behavior.

The divergence of ratings has also been researched. For example, Dimson et al. (2020) studied the correlations of MSCI’s, Sustainalytics’ and FTSE Russel’s ESG ratings; all three are prominent ESG providers. This can be looked at through evaluating the ”convergent validity” of the different ESG rater’s ESG scores. Although different providers may generate diverse scores due to measuring different aspects of ESG behavior, it should still be the case that evaluations or metrics that measure the same variable should generate positively correlated scores. Additionally, the score should not correlate with dissimilar or unrelated variables. However, by looking at 878 U.S. companies and comparing the overall ESG ratings assigned to them by MSCI and Sustainalytics in 2019, there was barely any perceptible relationship between the ratings, as illustrated in Figure 6:

LaBella et al. observe similar divergence between the MSCI and LSEG ESG ratings, as illustrated in Figure 7. Ultimately, the differences between the ESG raters are due to different views, weightings, and weighting bias of key indicators E, S, and G for various industries. In

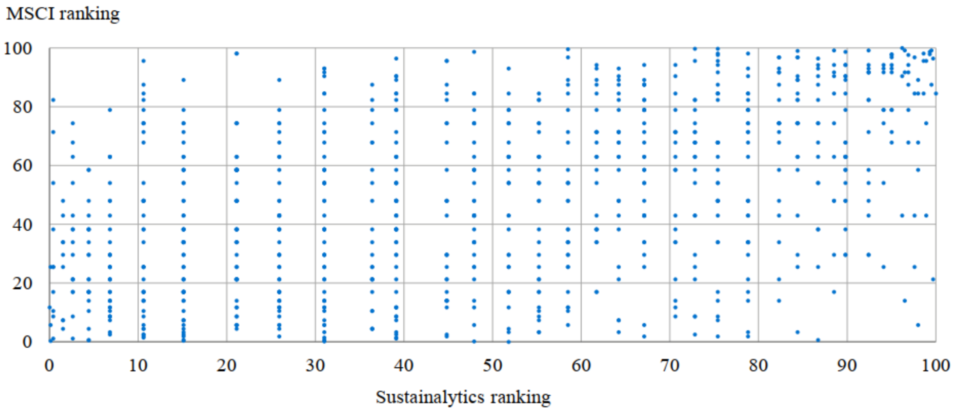

Figure 6. MSCI vs. Sustainalytics rankings at the beginning of 2019 (Dimson et al., 2020, 3).

addition, differences in methods and factors used to measure indicators inevitably also lead to discrepancies. These disparities are equally as dramatic both at the aggregate (company's combined ESG score, for example) and at the underlying ESG level. According to LaBella et al., this implies that the discrepancies are not only in how ESG factors are weighted, but also in the variation of factor definitions and metrics. An additional factor recognized by many authors is the fact that the ESG ratings of all providers generally include some amount of large cap and geographic bias, but also LaBella et al.. As rating agencies must rely on somewhat on the disclosure data of companies, this tends to consistently favor larger, multinational companies with more resources to produce sustainability disclosure documents, which leads to said companies getting better ESG ratings. Similarly, companies operating in jurisdictions with strict or clearer regulatory requirements for reporting on ESG practices tend to have better ESG ratings. To demonstrate this, Appendix 6.4 includes visualizations of these effects from LaBella et al.'s study.

An additional problem relating to using ESG data is the fact that despite there being multiple companies specializing partly or entirely in rating companies based on their commitment to ESG and ESG performance, these ratings are difficult to mix and match. For example, Puttonen and Puttonen (2021) write that despite multiple ESG rating agencies providing seemingly exact measurements, these vary from rating agency to rating agency. These differences are due to differences in methodologies for measuring ESG perfor-

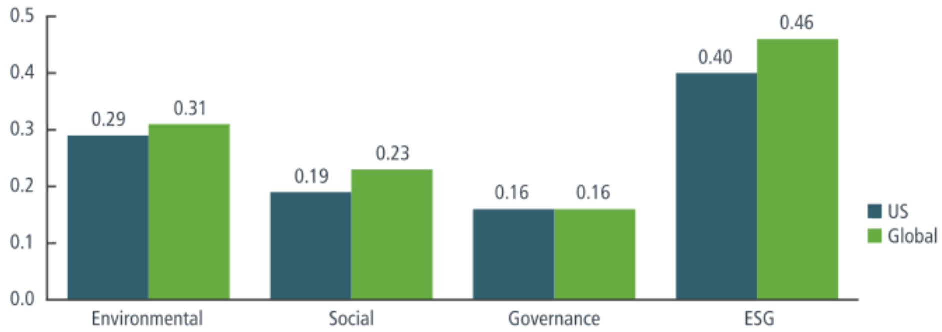

Figure 7. Historic Correlation of MSCI's and LSEG's ESG Ratings 2012-2018 (LaBella et al., 2019, 3).

mance and different definitions of sustainability and corporate responsibility. Therefore, despite there being at least 10 popular platforms that offer different ESG ratings for companies, their data could not be used to complement the data set due to lack of access and different methodologies.

As we have uncovered so far, there are several challenges related to the use of ESG data. A particular additional challenge related to ESG is that historically companies have not been required to report on, for example, their internal ESG-related initiatives. For example, in the US, ESG metrics are self-reported by the companies on the basis of ”material information”, which is defined by the U.S. Supreme Court as information, which is ”important to investors in making investment and voting decisions”. According to (Bender, Bridges, He, Lester, & Sun, 2018, 45) means that in the US companies do not need to comply with formal standards set by regulators, when deciding what ESG data, factors, and ESG-related information are material and should be disclosed to investors. In the EU, more formal ESG standards have been evolving, for example, the Accounting Directive of 2014 (*Directive 2013/34*, n.d.). This directive was part of the EU's wider Corporate Social Responsibility initiative, which includes plans for a consistent reporting approach to support smart, sustainable, and inclusive growth (Bender et al., 2018, 55).

5.1.3 Sustainalytics' ESG Data

Sustainalytics, founded in 2009 through the consolidation of several smaller ESG research firms has since become a leading provider of ESG ratings and research (Huber & Comstock, 2017). In 2020, Sustainalytics became a part of Morningstar Inc. Since 2017, its coverage has grown from over 6500 companies across 42 sectors to over 15000 companies globally in 2025 (Huber & Comstock, 2017; Sustainalytics, 2025b).

Sustainalytic's risk rating methodology is risk based and tries to assess a company's exposure to ESG risk and evaluate, how much of its ESG risk is "managed" or "unmanaged" (see Figure 8). According to Sustainalytics, the risk assessment process starts with considering a firm's total exposure to relevant ESG risks (outer yellow box) and then proceeds by qualifying, which part of that exposure is **manageable** vs. inherently **unamageable** given its industry (boxes 2 and 3). The manageable portion is affected by the company's management performance (box 4) and additionally any possible gaps or ESG controversies are considered *management gaps*. This leaves the unmanaged risk, which is the portion of risk not mitigated by the company. (Sustainalytics, 2024, 2025a)

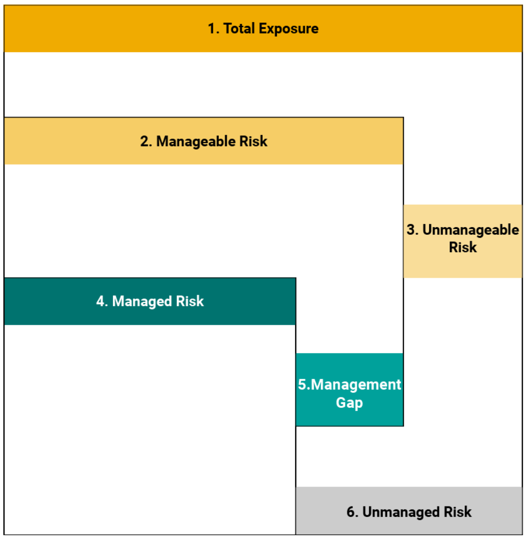

Figure 8. Sustainalytics Risk Framework (Sustainalytics, 2025a).

This unmanaged risk is what Sustainalytics considers the degree to which ESG issues could negatively impact the company's value and are not being addressed. Sustainalytics changed its ESG metrics in November 2019. Until then, the higher the ESG score meant that the more sustainable the company was. Since November 2019, Sustainalytics moved to a risk metric, where the lower the ESG score was, the less the company has ESG risk or the more sustainable the company is (Lienus10, 2021). To account for this, in this study, the score ”polarity” has been unified using the formula in 11, where the i refers to an ESG score at time i for data after November 2019.

$$\text{ESG score after 2019}_i = 100 - \text{ESG Score}_i \tag{11}$$

This makes the new methodology equivalent with the old one, where the higher scores meant better ESG performence.

5.1.4 LSEG data

Financial data for this study is collected from the LSEG database. The financial data collected consisted of monthly price data and market capitalizations. Share prices are collected to measure financial performance and returns on investment. The market capitalization data are used to divide the dataset into categories by ”size”. Data on company industries is also collected from LSEG using the Thomson Reuters Business Classification. In this study, companies in the financial services industry are not excluded.

Because several more recent studies on the ESG momentum (see, e.g. Chen & Yang, 2020; Koundouri, Landis, & Pittis, 2025; Magnani et al., 2024) have suggested that the ESG momentum may require more frequent rebalancing to capture abnormal returns, this study uses monthly rebalancing. For example, Galema and Gerritsen (2025), found that in the US stock market, MSCI ESG rating upgrades could be associated with abnormal outperformance over the next 6 months, particularly when environmental scores improved. These findings are in line with Nagy et al.'s (2016) observation that their ESG momentum strategy, which produced abnormal returns in the short term.

However, this is not possible using all ESG rating provider's data. For example, LSEG's ESG ratings are updated annually, which means that after lagging the ESG ratings to avoid look-ahead bias, the impact of ESG ratings changes on stock prices may have already evaporated. Therefore, in this study Sustainalytics' ESG scores are used sourced from

Yahoo Finance, as these are updated monthly. However, these were available mainly for US companies. For this purpose, I scraped ESG ratings for S&P500 companies from Yahoo Finance for the period 2013-2024. Yahoo uses Sustainalytic's overall ESG scores for the environmental, social and governance aspects of ESG, which are updated several times a year (Hale, 2016). Additionally, LSEG's ESG data seems to be sometimes misleading. An ESG assigned to a company for a given year does not suggest that that is its current ESG standing in LSEG's opinion. More often than not, LSEG will give a company an ESG rating for a given year i, but that score may be the score assigned to the company a number of years earlier, when the ESG rating was previously reviewed. This is an additional reason, which made using ESG scores from Sustainalytics the better option.

For the regressions, data from Kenneth R. French's (2021) database was used. These data consisted of yearly data on the Fama-French market factors: size, value, momentum, profitability, investment, and the risk-free rate of return. The reason for using Kenneth R. French's data set instead of replicating it using raw data is to achieve better comparability between other factor studies. Additionally, replicating the data set collected by Fama and French from the Compustat and WRDS databases is difficult if not impossible using an alternative data source like LSEG, which was the only one to which I had access.

5.2 Methodology

This subsection focuses on the methodology used in this study. First, portfolio construction is covered and relevant portfolio statistics presented. Finally, the methods and results of the study are discussed.

5.2.1 Portfolio construction

Earlier studies on momentum-related anomalies have suggested it is important to set a suitable threshold for including (excluding) companies into (from) momentum portfolios. Titman and Jegadeesh (1993) recognized this in their study, where they found that using top and bottom decile stocks for long-short portfolios produced better results.

The empirical design of this study follows one laid out by Magnani et. al (2024). The dataset being studied is members of the S&P 500 index between 2014-2024. To construct this dataset, LSEG's point-in-time data on S&P 500 membership is used to reconstruct monthly joiner/leaver logs from the S&P 500 index. The purpose of this step is to mitigate survivorship bias. For each month t, the constituents of the investable universe are therefore those companies, which have been observable at month end $t-1$. Figure 5 seen earlier illustrated the joiners and leavers from the S&P 500 index between 2014-2024.

As mentioned earlier, prices, market capitalisation (lagged for weighting), and sector classification (TRBC Sector Classification) data are sourced from LSEG/Refinitiv. Monthly total returns are used for calculating the returns and these include cash dividends, when available; when only close prices are available adjusted close prices have been used. ESG data (Total, E, S, G) are Sustainalytics' scores at a monthly frequency. Records are de-duplicated, aligned to month-ends, and lightly winsorised at the 1st-99th percentiles cross-sectionally each month. All signals are also lagged by one month to prevent look-ahead biasing the results.

For records to be eligible to be used in the analysis, they must, in addition to being present at month end $t-1$, have at least H months of valid ESG scores, which are required compute momentum ($H \in \{1, 3, 6, 12\}$) and W months to compute stability ($W \in 18, 24, 30, 36$). To ensure sector neutrality, sector buckets with less than 8 names are dropped. The following section (see 5.2.2) explains more thoroughly, how ESG scores are neutralized for size and industry biases prior to being used for computing ESG momentum or ESG volatility factors.

Using a range of periods for ESG momentum computation reflects the finding in momentum-related literature that the choice of period often impacts the result of empirical tests and thus adds robustness (see, e.g., Jegadeesh & Titman, 1993).

5.2.2 ESG score neutralization

In this study, ESG scores have been first put on a time-consistent scale as mentioned previously in section 5.1.3. This, again, was necessitated by the fact that this study uses

Sustainalytics' ESG scores, which underwent a change in rating methodology in 2019-11-01.

Earlier studies by, for example, Bruno et al. (2022) and Kaiser (2020) have established that ESG scores tend to be impacted by both company size (typically measured with market capilization) and vary significantly between different industries. Magnani et al. also observe these effects in their study and, therefore, develop a methodology to mitigate against such unwanted biases by rank-neutralizing their ESG scores by size and industry. This methodology is also implemented in this study.

First, ESG scores are corrected for possible size-related biases by dividing the stocks into five size-quantiles (size is measured by market capilization). For each quantile, the median ESG score is then determined and each individual ESG score subtracted by said median within each this quantile. This results in a new set of ESG scores, which are deviations from the size-specific medians and, therefore, size-neutral. These residuals are then converted into percentile ranks with values between 0-1 with 0 assigned to the companies with the highest negative deviation from the median and 1 to the companies with the highest positive deviation. Secondly, these now size-neutralized ESG scores are further sector-neutralized by applying a similar process, where the ESG scores are converted again into percentile ranks in $[0, 1]$ within each industry classification. This yields ESG scores that are rank-neutralised by size and industry and can then be compared across time and sectors.

5.2.3 Portfolio construction

Magnani et al. form two types of portfolios: ESG momentum and ESG volatility portfolios. The authors define ESG momentum as the percentage increase in a security's ESG over a certain period time. To create ESG momentum portfolios, the stocks in month x are first ordered in descending order by their ESG momentum, which is calculated using the previously unbiased ESG scores. Equally-weighted quintile portfolios are then formed based on this descending order resulting in the first quintile (Q1) containing those firms with the highest ESG momentum and Q5 those with the lowest ESG momentum. From these quintiles, a long-short portfolio is created. This ESG-MOM portfolio takes a long position in the quintile of companies, whose rank-neutralised ESG score has improved the most over the look-back period of H months and shorts the lowest momentum companies (companies whose ESG score deteriorated the most or improved the least).

$$\text{MOM portfolio return} = \text{Return}_{Q5} - \text{Return}_{Q1} \tag{12}$$

The ESG volatility portfolios are created in a fairly similar fashion. First, for each month x the ESG volatility is the sample standard deviation of the size- and industry-neutralized ESG scores over a fixed time window of W months, where $W \in \{18, 24, 30, 36\}$. Stocks are sorted in ascending order by their ESG volatility and equally-weighted quintile portfolios formed. The quintiles are such that Q1 contains the companies with the least volatile ESG scores and Q5 the most volatile over the previous W months.

$$r_t^{\text{VOL}} = r_t^{Q1} - r_t^{Q5}, \tag{13}$$

The main idea behind Magnani et al.'s research is that investors will reward ESG stability more than short-term ESG improvements. The authors define ESG momentum as the percentage increase in a security's ESG over a certain period time. Different ESG momentum and ESG volatility computation periods are used, as it's been established in the research literature that the results of momentum-related research can vary depending on the choice computation period. Thus, for a given month x, ESG momentum is computed using a range of computation periods, where $x \in \{1, 3, 6, 12\}$. Partially, using a range of computation periods in computing ESG momentum tries to account for the delay often present in ESG rating changes being communicated by ratings agencies. Hence, using a range of periods should also neutralise for these lags in rating changes being publicized.

The ESG momentum (ESG-MOM) portfolios are constructed by sorting the universe of stocks in descending order by their ESG momentum in any given month x. This ordered set of stocks are then divided into quintile portfolios that are equally weighted. The first quintile (Q1) will then consist of those stocks with the highest ESG momentum and the last (Q5) those with the lowest ESG momentum.

Similarly, ESG volatility, both in this and Magnani et al.'s study, is defined for month x as the sample standard deviation of size- and industry- by using a range of the six-month percentage change with a one-month skip (6-minus-1) and ESG stability is defined as the rolling volatility or the standard deviation of the neutralised ESG scores over 18-24 months.

Two portfolios are tested in this study. A portfolio using ESG-MOM (top-minus-bottom momentum quintiles, i.e., improvers - decliners) and ESG-VOL (low-volatility-minus-high-volatility of ESG scores, i.e., stable - unstable). The time-series performance of these portfolios are evaluated using the Fama-French 5-factors model and for robustness price momentum (UMD) and low volatility factor are included as well. To verify the results, a GRS test is ran on the results of the factor models. Performance analysis is also included by computing Sharpe ratios and mean returns for the different portfolios.

6 Results

This chapter presents the results of the portfolios constructed using the methodology outlined in section 5.2. The empirical results are analysed with the help of three different regression models: the Fama–French five-factor model (FF5), FF5 augmented with Momentum (Carhart, 1997), and FF5 augmented with the Low Volatility factor introduced by Blitz and Van Vliet (2007). These models are useful as they reflect the main systematic risk factors that explain stock returns, and following Magnani et al., I test whether ESG momentum or volatility strategies "earn returns beyond what is explained by these risk factors." The portfolio analysis using factor models is covered in section 6.2, with robustness testing in subsection 6.2.1. This section begins with descriptive statistics to motivate the subsequent sorting and neutralisation choices.

6.1 Descriptive statistics

As mentioned earlier, it has been established in the literature that ESG scores are significantly affected by firm size and that there are substantial differences in ESG scores between industries (see, e.g. Bruno et al., 2022; Kaiser, 2020). This phenomenon is also observed in the S&P 500 dataset.

Table 3 presents a size-sorted panel of stock returns and ESG scores. In it, clear systematic relations can be observed among size, returns, and ESG levels. Average monthly returns rise from 0.27% in Q1 to 1.08% in the two largest quantiles (Q4 and Q5). At the same time, return volatility falls monotonically from 5.25% to 3.59%. In lockstep with falling volatility, the Sharpe ratio also increases with size. Within this large-cap universe and sample window, a small-cap premium does not seem observable as larger firms appear to deliver higher risk-adjusted returns.

Similarly, total ESG scores rise from 59 in Q1 to 66 in Q5, with medians following the same pattern and with the dispersion of ESG scores (σ) being modestly lower for large companies. This mirrors the pattern reported by Magnani et al. and justifies applying their size neutralization before constructing ESG portfolios so as not to confound any ESG effect with an underlying size tilt.

Table 3. Summary statistics and ESG scores for five size-sorted quantiles (adapted from Magnani et al., 2024, 672).

	Stock returns				Sustainalytics ESG scores				
Size qtl	Mean (%)	σ (%)	Sharpe Ratio	Mark. Cap.	Mean	Median	σ	Max	Min
1	0.27	5.25	0.05	9,499	59.17	56.41	12.09	92.92	18
2	0.88	4.09	0.21	16,156	60.46	58.22	12.12	93.16	19.16
3	0.89	4.01	0.22	26,814	62.92	62.95	11.54	93.01	12.59
4	1.08	3.64	0.30	47,887	63.46	62.73	10.75	90.87	19.61
5	1.08	3.59	0.30	209,740	65.79	65.35	10.09	91.49	14.12

The industry summary (Table 4) also confirms strong sectoral heterogeneity in both returns and ESG. For example, Information Technology delivers the highest average monthly return ($\approx 1.66\%$/month) at the cost of relatively high volatility ($\approx 9.1\%$/month), while Utilities and Consumer Staples exhibit lower volatility (≈ 6–6.6%) and correspondingly lower mean returns ($\approx 0.6\%$/month). Energy industry's Sharpe ratio is weak owing to elevated volatility, which is consistent with inherent sector cyclicality.

Table 4. Summary statistics of returns and ESG scores for industries.

	Returns				Sustainalytics ESG scores				
Industry	Mean (%)	σ (%)	Sharpe Ratio	Mark. Cap.	Mean	Median	σ	Max	Min
Communication Serv.	0.89	8.31	0.11	179,708	65.53	60.78	13.14	89.68	26.08
Consumer Discret.	1.10	10.14	0.11	60,676	61.06	59	12.44	92.92	27.37
Consumer Stapl.	0.60	6.63	0.09	67,372	64.26	64.63	10.10	88.20	24.61
Energy	0.63	13.15	0.05	61,465	60.78	60.86	7.54	81.69	22
Financials	1.03	7.65	0.14	57,500	60.17	57.54	11.54	89.50	24.63
Health Care	1.04	7.86	0.13	71,853	61.65	60.91	10.87	89.10	28.99
Industrials	1.13	7.93	0.14	37,078	61.15	60.11	10.61	90.74	21.74
Information Technology	1.66	9.08	0.18	114,604	67.08	68.47	13.01	92.40	12.59
Materials	1.02	9.49	0.11	22,451	63.31	65	9.46	89.19	19.95
Real Estate	0.60	7.34	0.08	24,915	61.72	59	14.56	93.16	19.61
Utilities	0.60	6.01	0.10	27,927	65.22	65.96	7.95	81.38	19.13

Notes: Mean and σ are monthly; Mark. Cap. is the time-average of cross-sectional mean market cap (millions). ESG statistics use polarity-unified scores.

ESG levels vary substantially between industries: Information Technology and Communication Services are at the high end (means ≈ 67 and 65–66), whereas Financials and Energy are lower (≈ 60–61). Dispersion also differs: Real Estate and Financials have relatively wide spreads ($\sigma \approx 11.5$–14), suggesting greater between-firm differentiation, while Energy is tightly clustered ($\sigma \approx 7.5$). Maximum scores are similar across industries (high 80s/low 90s), but minima vary widely (from the high 20s in Energy to the high 50s in Utilities), indicating a sector-specific left tail.

Taken together, these observations justify two design choices applied here and in Magnani et al.: (i) rank-neutralisation by size followed by sector ranks to isolate any ESG signal from cross-sectional co-movements, and (ii) sector-neutral, value-weighted (lagged market cap) portfolio sorts to avoid inadvertent loadings from industry composition or size. The following section details the portfolio tests.

6.2 Portfolio analysis and factor models

This section reports results from regressions under multiple specifications for both ESG-VOL and ESG-MOM. Testing across look-backs (MOM) and volatility windows (VOL) reduces the chance that outcomes are specific to one arbitrary parameter choice and helps separate true signal from noise. The goal is to verify whether any abnormal returns are attributable to ESG rather than to known premia (market, size, value, profitability, investment, price momentum, defensive/low-volatility).

Before the factor tests, we characterise downside risk. Figure 9 compares the maximum drawdowns of ESG-VOL, ESG-MOM, and the S&P 500 from 2016 onward (the first date with sufficient data to form ESG-VOL). ESG-VOL exhibits comparatively shallow drawdowns, consistent with its near-market-neutral construction and lower payoff volatility. Figure 10 shows the path of drawdowns from prior peaks, reinforcing the same conclusion.

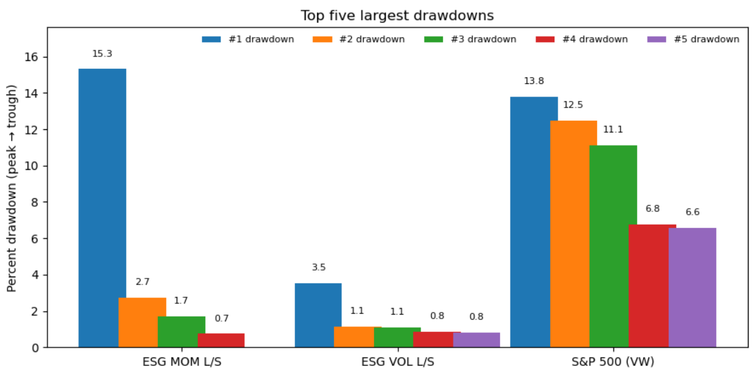

Figure 9. Top drawdowns.

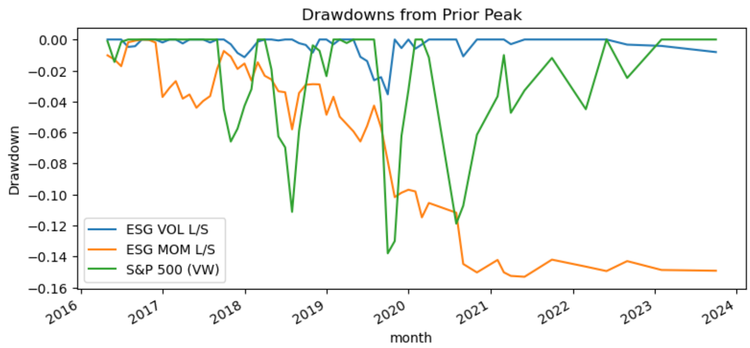

Figure 10. Drawdowns from prior peak (path).

As a baseline control, the FF5 model is used to test whether either ESG strategy simply repackages broad exposures such as profitability (RMW) or investment (CMA), and whether the resulting returns are orthogonal to these factors. Price momentum (UMD) is added to check whether alpha remains after removing pure price momentum, and a defensive/low-volatility factor (DEF) is included to test whether ESG stability adds information beyond a standard low-vol tilt. An economically large and statistically significant intercept across these variants would indicate returns not explained by conventional premia.

Across look-back horizons, ESG-MOM delivers different outcomes (Table 5). At short horizons ($H{=}1, 3$), average returns and alphas are near zero with insignificant Newey–West t-statistics. At $H{=}6$, the factor turns decisively negative (mean -0.35%/month; annual mean -4.16%; annualised Sharpe ≈ -1.06) with a statistically significant alpha (-0.45%/month, $t_\alpha{=}-5.06$). Extending to $H{=}12$ reduces the magnitude and restores insignificance. Overall, ESG-MOM appears to reverse at intermediate horizons, while very short and longer horizons do not yield robust abnormal returns.

Across volatility windows $W \in \{18, 24, 30, 36\}$ months, ESG-VOL delivers economically meaningful returns, strongest at intermediate windows (Table 6). At $W{=}24$, the strategy earns 0.48%/month ($\approx 5.74\%$/year) with a Newey–West t on alpha of 3.44 and an

Table 5. ESG Momentum (long–short) results across look-back horizons H.

H	Mean (%, m)	t_{NW}	Ann. mean (%, y)	Ann. Sharpe	N	Alpha (%, m)	t_α	Alpha (%, y)
1	-0.00	-0.0035	-0.00	-0.001	77	0.09	1.0739	1.07
3	0.01	0.1152	0.12	0.037	75	0.02	0.3318	0.30
6	-0.35	-3.4726	-4.16	-1.064	72	-0.45	-5.0561	-5.34
12	0.02	0.1890	0.28	0.073	66	-0.11	-1.0767	-1.35

Notes: Monthly and annual returns are expressed in percent. Newey–West t-statistics (t_{NW} for the mean and t_α for the factor alpha) use 6 lags. Ann. Sharpe is the annualized Sharpe ratio of the monthly long–short series.

annualised Sharpe of 1.40. Windows of 18 and 30 months also show significant alphas (NW$_t$ ≈ 3.22 and 2.82) and healthy Sharpe ratios (≈ 1.25). Performance attenuates at $W{=}36$, where the signal weakens. This suggests that the "stability premium" is most pronounced when ESG volatility is measured over roughly 1.5–2.5 years and that very long windows dilute the signal.

Table 6. ESG Stability (VOL) long–short results across volatility windows W.

W	Mean (%, m)	t_{NW}	Ann. mean (%, y)	Ann. Sharpe	N	Alpha (%, m)	t_α	Alpha (%, y)
18	0.37	3.466	4.44	1.255	61	0.40	3.222	4.84
24	0.44	3.123	5.32	1.404	55	0.48	3.435	5.74
30	0.28	2.796	3.36	1.247	49	0.30	2.818	3.57
36	0.11	0.865	1.38	0.477	43	0.13	1.031	1.61

Notes: Monthly and annual returns are expressed in percent. Newey–West t-statistics (t_{NW} for the mean and t_α for the regression alpha) use the lag length specified in the main text. Ann. Sharpe is computed from monthly long–short returns. Alphas are estimated from factor regressions consistent with the paper's baseline.

Regressing the ESG-VOL portfolio's monthly excess returns on FF5 plus UMD and DEF, the intercept (alpha) measures abnormal return unexplained by these premia. Using Newey–West standard errors, we find an alpha of 0.40%/month at $W{=}18$ ($t_{\text{HAC}}{=}3.222$, $p{=}0.001$), corresponding to about 4.84%/year. This indicates the performance is not explained by conventional factors. (Full regressions appear in Appendix 6.4.)

The null hypothesis $H_0 : \alpha = 0$ posits that ESG-VOL returns are fully explained by factor loadings and that adding ESG-VOL brings no incremental value. Because $\alpha = 0.40\%$/month is statistically significant ($z = 3.22, p = 0.001$; 95% CI $[0.002, 0.006]$/month) and other loadings are insignificant, therefore H_0 is rejected.

Table 7. ESG volatility portfolio regression with $W = 24$.

	y
const	0.004***
	(0.001)
Mkt-RF	-0.009
	(0.015)
SMB	0.072
	(0.066)
HML	-0.024
	(0.038)
RMW	-0.015
	(0.082)
CMA	-0.004
	(0.055)
UMD	0.017
	(0.031)
DEF	-0.027
	(0.064)
R-squared	0.035
R-squared Adj.	-0.153

Under CAPM, ESG-VOL's alpha is 0.27%/month (t_{NW}=3.14), implying 3.16%/year (simple) and 3.31%/year (compounded) of excess return of what a market-neutral strategy should earn. The CAPM beta is essentially zero ($\beta \approx -0.006$), and Jensen's alpha equals the CAPM alpha over 61 months, indicating market-neutral abnormal performance. Therefore, the long-stable minus short-unstable portfolio delivers returns not captured by standard factors and improves cross-sectional pricing. Table 7 summarises the W=24 regression.

Figure 11 plots cumulative wealth for ESG-VOL and the S&P 500. Despite lagging the index in bull markets, ESG-VOL's near-zero beta and positive alpha make it a potentially valuable overlay.

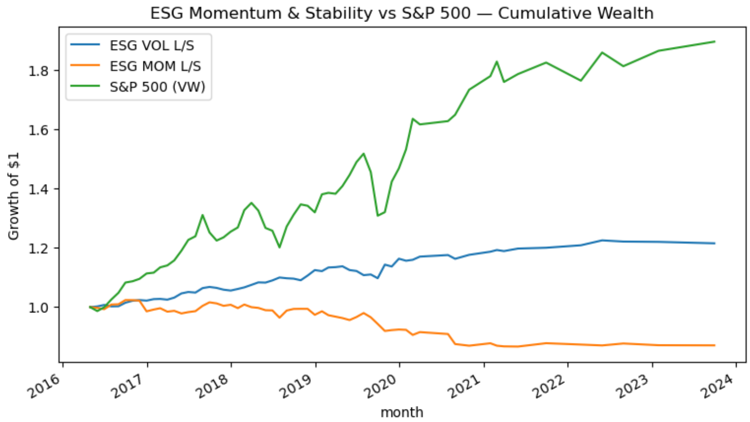

Figure 11. Cumulative wealth of different strategies and S&P 500 (2016-2024).

To ensure robustness, we next run a GRS test on the ESG-VOL results.

6.2.1 Robustness testing

We implement the Gibbons, Ross, and Shanken (1989) test to assess whether ESG-MOM and ESG-VOL add pricing power. The GRS joint null is $H_0 : \alpha_1 = \cdots = \alpha_N = 0$. This guards against cherry-picking and evaluates whether a factor improves the joint pricing of multiple test portfolios. Here, I compare FF5, FF5+ESG-Vol, FF5+UMD, and FF5+UMD+ESG-Vol. Improvement is indicated by a lower GRS statistic and higher p-value. Results appear in Table 8.

Table 8. GRS tests across factor sets (test portfolios: 5 quintiles; monthly data).

Model	K	T	GRS	p-value	df$_2$
FF5	5	55	2.22	0.069	45
FF5 + UMD	6	55	2.11	0.083	44
FF5 + ESG-Vol	6	55	1.69	0.157	44
FF5 + UMD + ESG-Vol	7	55	1.61	0.177	43
FF5 + UMD + DEF + ESG-Vol	8	44	1.11	0.375	31

K = number of risk factors in the model excluding the intercept (e.g., FF5: $K = 5$, after adding UMD $K = 6$. N = number of test portfolios (here, $N = 5$, the quintiles). T = The number of time observations (here, months), after aligning returns and factors. df$_2 = T - N - K$, the second degree of freedom.

The GRS statistics indicate that plain FF5 only marginally prices the five test portfolios:

GRS$= 2.22$ with $p = 0.069$ (df$= (5, 45)$) implies borderline rejection of the joint-alpha null at the 10% level. Adding UMD yields a modest improvement (GRS$= 2.11, p = 0.083$). By contrast, adding ESG-Vol materially enhances cross-sectional fit: FF5+ESG-Vol reduces the statistic to 1.69 and raises p to 0.157 (df$= (5, 44)$), so we fail to reject at conventional levels—consistent with ESG stability providing incremental explanatory power. Combining UMD and ESG-Vol further improves fit (GRS$= 1.61$, $p = 0.177$). The strongest result arises when also including DEF (GRS$= 1.11$, $p = 0.375$, df$= (5, 31)$), despite fewer observations due to factor availability. Overall, the monotonic improvement as ESG-Vol and DEF adds further support to the interpretation that ESG stability is a distinct, economically meaningful dimension.

6.3 Overall Discussion

The evidence suggests that, with the data and portfolio construction used here, ESG momentum does not robustly appear as a priced risk factor. While point estimates are sometimes positive, they are statistically fragile and sensitive to the treatment of serial correlation and the look-back/skip conventions. Under a conservative HAC/Newey–West setting with a 12-month lag for monthly data, the signal is at best borderline.

By contrast, the ESG stability factor delivers economically meaningful and statistically significant alphas in most specifications, including FF5 and FF5+UMD+LowVol. Notably, factor loadings on the conventional premia are small and mostly insignificant, suggesting ESG-VOL is not a repackaging of those premia. From a portfolio perspective, the strategy's near-zero beta and low correlation with the market mean that — even though cumulative wealth produced by the strategy lags the index in (see figure 11)- its risk-adjusted contribution remains positive and potentially valuable as an overlay.

Why might stability show up while momentum does not? Measurement error and methodology changes in ESG ratings can attenuate short-horizon signals more than long-horizon or dispersion-based ones. Vendors often update scores in lumps rather than continuously, creating "bunching" and staleness, which particularly affects LSEG data. This can dilute month-to-month momentum yet leave cross-sectional volatility relatively intact.

Economically, stability may proxy for lower non-financial risk or stronger organisational processes, reducing left-tail outcomes and possibly discount rates, whereas ESG momentum—driven by slower-moving disclosures—may already be incorporated by investors.

Turnover for stability is moderate but non-trivial; after trading costs and short-rebate assumptions, net alphas would be lower. Nonetheless, because ESG-VOL is close to market-neutral, it can serve as a diversifying overlay to a long-only core without materially increasing market exposure.

6.4 Limitations and Further Research

The limitations of this study are fairly clear. Firstly, the universe being the S&P 500 constituents with survivorship and membership dynamics approximated from available data means that in different universes including more middle or small cap firms or firms in a different region could all yield different conclusions. Secondly, the construction relies on rank-neutralization by sector and size. Different neutralisation methods might change the cross-section and results. This study also abstracts away transaction costs, borrow fees, and financing of the short leg, which could all materially impact the performance of the portfolios.

Future research should examine out-of-sample universes and incorporate transaction cost models with short-rebate assumptions to verify these results. Exploring alternative stability definitions could also be valuable to further examine the interactions between ESG stability and financial momentum and quality. These could help determine whether the stability premium is complementary or, for example, substitutable.

Future research should also incorporate alternative ESG data providers or combine multiple sources to reduce measurement error. A deeper exploration of industry-specific dynamics might also be useful. For example, in sectors heavily scrutinized for environmental or social impacts, ESG improvements might convey more material information than in sectors where such issues are less prominent. Finally, longer historical series (should they become available) would allow for more robust tests. Richer microdata on event-timing of ESG controversies could also help disentangle whether stability carries a risk-based interpretation. If stability is risk-based, the firms with stable ESG scores should suffer fewer and less severe negative events.

Overall, the thesis adds to the evidence that ESG stability seems to behave like a diversifying, alpha-generating overlay, at least, in large-cap U.S. equities over the studied sample period. The evidence for ESG momentum remains inconclusive.

Bibliography

Alves, R., Krueger, P., & van Dijk, M. A. (2023, December). *Drawing Up the Bill: Are ESG Ratings Related to Stock Returns Around the World?* (SSRN Scholarly Paper No. 25-41). Rochester, NY: Social Science Research Network. https://doi.org/10.2139/ssrn.4674146

Amihud, Y., & Mendelson, H. (1986, December). Asset pricing and the bid-ask spread. *Journal of Financial Economics*, *17*(2), 223–249. (Number: 2) https://doi.org/10.1016/0304-405X(86)90065-6

Bannier, C. E., Bofinger, Y., & Rock, B. (2019). *Doing safe by doing good: ESG investing and corporate social responsibility in the U.S. and Europe* (Working Paper No. 621). CFS Working Paper Series. (Issue: 621) Retrieved from [2023-11-25]https://www.econstor.eu/handle/10419/200161

Banz, R. W. (1981, March). The relationship between return and market value of common stocks. *Journal of Financial Economics*, *9*(1), 3–18. (Number: 1) https://doi.org/10.1016/0304-405X(81)90018-0

Basu, S. (1977). Investment Performance of Common Stocks in Relation to Their Price-Earnings Ratios: A Test of the Efficient Market Hypothesis. *The Journal of Finance*, *32*(3), 663–682. (Number: 3 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.1977.tb01979.x) https://doi.org/10.1111/j.1540-6261.1977.tb01979.x

Bauer, B., Koedijk, K., & Otten, R. (2002, August). International Evidence on Ethical Mutual Fund Performance and Investment Style. *Journal of Banking & Finance*, *29*.

Bender, J., Bridges, T., He, C., Lester, A., & Sun, X. (2018). A blueprint for integrating ESG into equity portfolios. *Journal of investment management : JOIM*, *16*(1).

Berg, F., Kölbel, J. F., & Rigobon, R. (2022, November). Aggregate Confusion: The Divergence of ESG Ratings*. *Review of Finance*, *26*(6), 1315–1344. https://doi.org/10.1093/rof/rfac033

Black, F. (1972). Capital Market Equilibrium with Restricted Borrowing. *The Journal of Business*, *45*(3), 444–455. (Number: 3 Publisher: University of Chicago Press) Retrieved from [2024-05-09]https://www.jstor.org/stable/2351499

Black, F. (1993, October). Beta and Return. *The Journal of Portfolio Management*, *20*(1), 8–18. (Number: 1 Company: Institutional Investor Journals Distributor: Institutional Investor Journals Institution: Institutional Investor Journals Label: Institutional Investor Journals Publisher: Portfolio Management Research Section: Primary Article) https://doi.org/10.3905/jpm.1993.409462

Blitz, D. C., & Vliet, P. v. (2007, October). The Volatility Effect. *The Journal of Portfolio Management*, *34*(1), 102–113. (Company: Institutional Investor Journals Distributor: Institutional Investor Journals Institution: Institutional Investor Journals Label: Institutional Investor Journals Publisher: Portfolio Management Research Section: Asset Classes) https://doi.org/10.3905/jpm.2007.698039

Blowfield, M., & Murray, A. (2008). *Corporate Responsibility: A Critical Introduction*. OUP Oxford. (Google-Books-ID: maX7ngEACAAJ)

Blume, M. E. (1975). Betas and Their Regression Tendencies. *The Journal of Finance*, *30*(3), 785–795. (Number: 3 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.1975.tb01850.x) https://doi.org/10.1111/j.1540-6261.1975.tb01850.x

Bodie, Z., Kane, A., Marcus, A. J., Switzer, L., Stapleton, M., Boyko, D., & Panasian, C. (2019). *Investments*. McGraw-Hill Education. (Google-Books-ID: n9YxxQEACAAJ)

Boffo, R., & Patalano, R. (2020). *ESG Investing: Practices, Progress and Challenges* (Tech. Rep.). Paris: OECD. Retrieved from https://www.oecd.org/finance/ESG-Investing-Practices-Progress-Challenges.pdf

Bondt, W. F. M. D., & Thaler, R. (1985). Does the Stock Market Overreact? *The Journal of Finance*, *40*(3), 793–805. (Publisher: [American Finance Association, Wiley]) https://doi.org/10.2307/2327804

Bruno, G., Esakia, M., & Goltz, F. (2022, March). "Honey, I Shrunk the ESG Alpha": Risk-Adjusting ESG Portfolio Returns. *The Journal of Investing*, *31*(3), 45–61. (Number: 3) https://doi.org/10.3905/joi.2021.1.215

Brzeszczyński, J., & McIntosh, G. (2014, March). Performance of Portfolios Composed of British SRI Stocks. *Journal of Business Ethics*, *120*(3), 335–362. https://doi.org/10.1007/s10551-012-1541-x

Busch, T., Bauer, R., & Orlitzky, M. (2016, March). Sustainable Development and Financial Markets: Old Paths and New Avenues. *Business & Society*, *55*(3), 303–329. (Number: 3 Publisher: SAGE Publications Inc) https://doi.org/10.1177/0007650315570701

Campbell, J. Y., & Cochrane, J. H. (1999, July). *Explaining the Poor Performance of Consumption-Based Asset Pricing Models* [Working Paper]. National Bureau of Economic Research. (Issue: 7237) https://doi.org/10.3386/w7237

Campbell, J. Y., & Vuolteenaho, T. (2004). Bad Beta, Good Beta. *The American Economic Review*, *94*(5), 1249–1275. (Number: 5 Publisher: American Economic Association) Retrieved from [2024-04-27]https://www.jstor.org/stable/3592822

Caplan, L., Griswold, J. S., & Jarvis, W. F. (2013, September). *From SRI to ESG: The Changing World of Responsible Investment* (Tech. Rep.). Common Fund Institute. Retrieved from https://files.eric.ed.gov/fulltext/ED559300.pdf

Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. *The Journal of Finance*, *52*(1), 57–82. (Number: 1 Publisher: [American Finance Association, Wiley]) https://doi.org/10.2307/2329556

Chen, H.-Y., & Yang, S. S. (2020, October). Do Investors exaggerate corporate ESG information? Evidence of the ESG momentum effect in the Taiwanese market. *Pacific-Basin Finance Journal*, *63*, 101407. https://doi.org/10.1016/j.pacfin.2020.101407

Chopra, N., Lakonishok, J., & Ritter, J. R. (1992, April). Measuring abnormal performance: Do stocks overreact? *Journal of Financial Economics*, *31*(2), 235–268. (Number: 2) https://doi.org/10.1016/0304-405X(92)90005-I

Conrad, J., & Kaul, G. (1988). Time-Variation in Expected Returns. *The Journal of Business*, *61*(4), 409–425. (Number: 4 Publisher: University of Chicago Press) Retrieved from [2024-04-07]https://www.jstor.org/stable/2352789

(Deceased), J., & C, M. (1967, May). *The Performance of Mutual Funds in the Period 1945-1964* [SSRN Scholarly Paper]. Rochester, NY: Social Science Research Network. https://doi.org/10.2139/ssrn.244153

De Lucia, C., Pazienza, P., & Bartlett, M. (2020, January). Does Good ESG Lead to Better Financial Performances by Firms? Machine Learning and Logistic Regression Models of Public Enterprises in Europe. *Sustainability*, *12*(13), 5317. (Number: 13 Publisher:

Multidisciplinary Digital Publishing Institute) https://doi.org/10.3390/su12135317

Dimson, E., Marsh, P., & Staunton, M. (2020, November). Divergent ESG Ratings. *The Journal of Portfolio Management*, *47*(1), 75–87. (Company: Institutional Investor Journals Distributor: Institutional Investor Journals Institution: Institutional Investor Journals Label: Institutional Investor Journals Publisher: Portfolio Management Research Section: Asset Classes) https://doi.org/10.3905/jpm.2020.1.175

Directive 2013/34. (n.d.). (titleTranslation: titleTranslation: titleTranslation:) Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:02013L0034-20240109

EUROSIF. (2016). *European SRI Study 2016* (Tech. Rep. No. 7). Brussels: Eurosif. Retrieved from [2025-06-09]https://www.eurosif.org/news/eurosif-report-2016/

EUROSIF. (2018). *European SRI Study 2018* (Tech. Rep.). Retrieved from [2025-03-15]https://www.eurosif.org/wp-content/uploads/2021/10/European-SRI-2018-Study.pdf

EUROSIF. (2021a). *Responsible Investment Strategies.* Retrieved from [2023-11-26]https://www.eurosif.org/responsible-investment-strategies/

EUROSIF. (2021b). *What are sustainable investments?* Retrieved from [2025-06-09]https://www.eurosif.org/sustainable-investments/

EUROSIF. (2022, December). *Sustainable investment.* Retrieved from [2022-12-06]https://www.eurosif.org/sustainable-investment/

Fama, E., & Kenneth, F. (2021). *Kenneth R. French - Data Library*. Retrieved from [2021-12-23]https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Fama, E. F. (1965). The Behavior of Stock-Market Prices. *The Journal of Business*, *38*(1), 34–105. (Number: 1 Publisher: University of Chicago Press) Retrieved from [2024-04-06]https://www.jstor.org/stable/2350752

Fama, E. F. (1991). Efficient Capital Markets: II. *The Journal of Finance*, *46*(5), 1575–1617. (Number: 5 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.1991.tb04636.x) https://doi.org/10.1111/j.1540-6261.1991.tb04636.x

Fama, E. F., & French, K. R. (1988). Permanent and Temporary Components of Stock Prices. *Journal of Political Economy*, *96*(2), 246–273. (Number: 2 Publisher: University of Chicago Press) Retrieved from [2024-04-07]https://www.jstor.org/stable/

1833108

Fama, E. F., & French, K. R. (1992). The Cross-Section of Expected Stock Returns. *The Journal of Finance*, *47*(2), 427–465. (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.1992.tb04398.x) https://doi.org/10.1111/j.1540-6261.1992.tb04398.x

Fama, E. F., & French, K. R. (1996). Multifactor Explanations of Asset Pricing Anomalies. *The Journal of Finance*, *51*(1), 55–84. (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.1996.tb05202.x) https://doi.org/10.1111/j.1540-6261.1996.tb05202.x

Fama, E. F., & French, K. R. (2006). The Value Premium and the CAPM. *The Journal of Finance*, *61*(5), 2163–2185. (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.2006.01054.x) https://doi.org/10.1111/j.1540-6261.2006.01054.x

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

Fama, E. F., & French, K. R. (2018, May). Choosing factors. *Journal of Financial Economics*, *128*(2), 234–252. https://doi.org/10.1016/j.jfineco.2018.02.012

Fama, E. F., & MacBeth, J. D. (1973). Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy*, *81*(3), 607–636. (Publisher: University of Chicago Press) Retrieved from [2024-05-09]https://www.jstor.org/stable/1831028

Fauzi, H. (2009, October). *Corporate Social and Financial Performance: Empirical Evidence from American Companies* [SSRN Scholarly Paper]. Rochester, NY. (Issue: 1489494) Retrieved from [2023-12-08]https://papers.ssrn.com/abstract=1489494

Fowler, S. J., & Hope, C. (2007, December). A Critical Review of Sustainable Business Indices and their Impact. *Journal of Business Ethics*, *76*(3), 243–252. https://doi.org/10.1007/s10551-007-9590-2

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

Galema, R., & Gerritsen, D. (2025, April). ESG rating changes and stock returns. *Journal of International Money and Finance*, *154*, 103309. https://doi.org/10.1016/j.jimonfin.2025.103309

Gibbons, M. R., Ross, S. A., & Shanken, J. (1989). A Test of the Efficiency of a Given Portfolio. *Econometrica*, *57*(5), 1121–1152. (Publisher: [Wiley, Econometric Society]) https://doi.org/10.2307/1913625

Giese, G., Lee, L.-E., Melas, D., Nagy, Z., & Nishikawa, L. (2019, July). Foundations of ESG Investing: How ESG Affects Equity Valuation, Risk, and Performance. *The Journal of Portfolio Management*, *45*(5), 69–83. https://doi.org/10.3905/jpm.2019.45.5.069

Giese, G., Nagy, Z., & Lee, L.-E. (2020, December). Deconstructing ESG Ratings Performance: Risk and Return for E, S, and G by Time Horizon, Sector, and Weighting. *The Journal of Portfolio Management*. https://doi.org/10.3905/jpm.2020.1.198

Global Sustainable Investment Alliance. (2013). *Global Sustainable Investment Review 2012* (Tech. Rep. No. 1). Global Investment Alliance. (Issue: 1) Retrieved from [2025-01-16]https://www.gsi-alliance.org/wp-content/uploads/2023/11/2012-Global-Sustainable-Investment-Review.pdf

Grobys, K. (2014). *Essays on empirical asset pricing* (doctoralThesis, Vaasan yliopisto). (Accepted: 2018-05-25T11:17:11Z ISBN: 9789524765756 ISSN: 2323-9123) Retrieved from [2024-05-12]https://osuva.uwasa.fi/handle/10024/7292

Hale, J. (2016, September). *How Sustainalytics Does Company ESG Research*. Retrieved from [2025-07-03]https://finance.yahoo.com/news/sustainalytics-does-company-esg-research-100000652.html

Hong, H., & Kacperczyk, M. (2009, July). The price of sin: The effects of social norms on markets. *Journal of Financial Economics*, *93*(1), 15–36. https://doi.org/10.1016/j.jfineco.2008.09.001

Huber, B. M., & Comstock, M. (2017, July). *ESG Reports and Ratings: What They Are, Why They Matter*. Retrieved from [2025-09-08]https://corpgov.law.harvard.edu/2017/07/27/esg-reports-and-ratings-what-they-are-why-they-matter/

IOSCO. (2021, November). *Recommendations on Sustainability-Related Practices, Policies, Procedures and Disclosure in Asset Management* (Tech. Rep. No. FR08/21). International Organization of Securities Commissions. (Issue: FR08/21) Retrieved from [2025-01-23]http://www.iosco.org/library/pubdocs/pdf/IOSCOPD688.pdf

Jaffe, J. F. (1974). Special Information and Insider Trading. *The Journal of Business*, *47*(3), 410–428. (Publisher: University of Chicago Press) Retrieved from [2024-03-25]https://www.jstor.org/stable/2352458

Jegadeesh, N., & Titman, S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *The Journal of Finance*, *48*(1), 65–91. (Publisher: [American Finance Association, Wiley]) https://doi.org/10.2307/2328882

Jensen, M. C., Black, F., & Scholes, M. S. (1972). *The Capital Asset Pricing Model: Some Empirical Tests* [SSRN Scholarly Paper]. Rochester, NY. Retrieved from [2024-05-06]https://papers.ssrn.com/abstract=908569

Kaiser, L. (2020, February). ESG integration: value, growth and momentum. *Journal of Asset Management*, *21*(1), 32–51. https://doi.org/10.1057/s41260-019-00148-y

Khan, M., Serafeim, G., & Yoon, A. (2016). Corporate Sustainability: First Evidence on Materiality. *The Accounting Review*, *91*(6), 1697–1724. (Publisher: American Accounting Association) Retrieved from [2023-12-03]https://www.jstor.org/stable/24907173

Kotsantonis, S., & Serafeim, G. (2019). Four Things No One Will Tell You About ESG Data. *Journal of Applied Corporate Finance*, *31*(2), 50–58. (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/jacf.12346) https://doi.org/10.1111/jacf.12346

Koundouri, P., & Landis, C. (2023). 6. ESG Momentum in International equity returns and the SDG content of financial asset portfolios. In *Transforming our world: Interdisciplinary insights on the sustainable development goals* (pp. 1–125). Paris: UN SDSN. (https://egd-report-2023.unsdsn.org/esg-momentum-in-international-equity-returns-and-the-sdg-content-of-financial-asset-portfolios/) Retrieved from [2025-07-03]https://egd-report-2023.unsdsn.org/esg-momentum-in-international-equity-returns-and-the-sdg-content-of-financial-asset-portfolios/

Koundouri, P., Landis, C., & Pittis, N. (2025, February). ESG Momentum in International Equity Returns and the SDG content of financial asset portfolios. *DEOS Working Papers*. (Number: 2523 Publisher: Athens University of Economics and Business) Retrieved from [2025-05-10]https://ideas.repec.org//p/aue/wpaper/2523.html

LaBella, M. J., Sullivan, L., Russell, J., & Novikov, D. (2019, September). *The Devil is in* *the Details: The Divergence of ESG Data and Implications for Responsible Investing*. QS Investors. Retrieved from [2025-06-24]https://www.firstlinks.com.au/uploads/Whitepapers/lm-qs-the-devil-is-in-the-details-0919.pdf

Lakonishok, J., & Shapiro, A. C. (1986). Systematic risk, total risk and size as determinants of stock market returns. *Journal of Banking & Finance*, *10*(1), 115–132. (Publisher: Elsevier) Retrieved from [2024-05-12]https://econpapers.repec.org/article/eeejbfina/v_3a10_3ay_3a1986_3ai_3a1_3ap_3a115-132.htm

Li, Q., Watts, E. M., & Zhu, C. (2024, April). *Retail Investors and ESG News* [SSRN Scholarly Paper]. Rochester, NY: Social Science Research Network. https://doi.org/10.2139/ssrn.4384675

Liang, H., & Renneboog, L. (2021, February). Corporate Social Responsibility and Sustainable Finance. In *Oxford Research Encyclopedia of Economics and Finance*. https://doi.org/10.1093/acrefore/9780190625979.013.592

Lienus10. (2021). *YESG - Yahoo ESG Scores for Python*. Retrieved from [2025-09-21]https://github.com/Lienus10/yesg/tree/master?tab=readme-ov-file#return-description

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. (Publisher: The MIT Press) https://doi.org/10.2307/1924119

Lo, A. W., & MacKinlay, A. C. (1988). Stock Market Prices do not Follow Random Walks: Evidence from a Simple Specification Test. *The Review of Financial Studies*, *1*(1), 41–66. (Publisher: [Oxford University Press, Society for Financial Studies]) Retrieved from [2024-04-07]https://www.jstor.org/stable/2962126

LSEG. (2025, March). *ESG Factor in Methodology*. LSEG Sovereign Sustainability Solutions. Retrieved from [2025-06-20]https://www.lseg.com/content/dam/data-analytics/en_us/documents/methodology/lseg-esg-factor-in-methodology.pdf

Magnani, M., Guidolin, M., & Berk, I. (2024, December). Strong vs. stable: the impact of ESG ratings momentum and their volatility on the cost of equity capital. *Journal of Asset Management*, *25*(7), 666–699. https://doi.org/10.1057/s41260-024-00377-w

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

91. (Number: 1 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.1952.tb01525.x) https://doi.org/10.1111/j.1540-6261.1952.tb01525.x

Martini, A. (2021, November). Socially responsible investing: from the ethical origins to the sustainable development framework of the European Union. *Environment, Development and Sustainability*, *23*(11), 16874–16890. (Number: 11) https://doi.org/10.1007/s10668-021-01375-3

Mossin, J. (1966). Equilibrium in a Capital Asset Market. *Econometrica*, *34*(4), 768–783. (Publisher: [Wiley, Econometric Society]) https://doi.org/10.2307/1910098

MSCI. (2025, March). *ESG Ratings | MSCI*. Retrieved from [2025-06-20]https://www.msci.com/data-and-analytics/sustainability-solutions/esg-ratings

Nagy, Z., Cogan, D., & Sinnreich, D. (2013, February). *Optimizing Environmental, Social and Governance Factors in Portfolio Construction: Analysis of Three ESG-Tilted Strategies* [SSRN Scholarly Paper]. Rochester, NY. https://doi.org/10.2139/ssrn.2221524

Nagy, Z., Kassam, A., & Lee, L.-E. (2016, May). Can ESG Add Alpha? An Analysis of ESG Tilt and Momentum Strategies. *The Journal of Investing*, *25*(2), 113–124. https://doi.org/10.3905/joi.2016.25.2.113

Nauman, B. (2019, September). Credit rating agencies join battle for ESG supremacy. *Financial Times*. Retrieved from [2025-03-15]https://archive.is/lPjaq

Novy-Marx, R. (2013, April). The other side of value: The gross profitability premium. *Journal of Financial Economics*, *108*(1), 1–28. https://doi.org/10.1016/j.jfineco.2013.01.003

OICU-ISCO. (2021, June). *Recommendations on Sustainability-Related Practices, Policies, Procedures and Disclosure in Asset Management - Consultation Report*. The Board of the International Organization of Securities Commissions. Retrieved from [2025-03-15]https://www.iosco.org/library/pubdocs/pdf/IOSCOPD679.pdf

Puttonen, T., & Puttonen, V. (2021). *Vastuullinen sijoittaminen teoriassa ja käytännössä*. Aalto University. (Accepted: 2021-09-30T09:00:07Z ISSN: 1799-4802 (electronic)) Retrieved from [2022-03-27]https://aaltodoc.aalto.fi:443/handle/123456789/110203

Reinganum, M. R. (1981). A New Empirical Perspective on the CAPM. *The Journal of*

Financial and Quantitative Analysis, *16*(4), 439–462. (Publisher: Cambridge University Press) https://doi.org/10.2307/2330365

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

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

Schueth, S. (2003). Socially Responsible Investing in the United States. *Journal of Business Ethics*, *43*(3), 189–194. (Publisher: Springer) Retrieved from [2024-03-27]https://www.jstor.org/stable/25074988

Schwert, G. W. (2002, October). *Anomalies and Market Efficiency* [SSRN Scholarly Paper]. Rochester, NY. https://doi.org/10.2139/ssrn.338080

Seyhun, H. N. (1986, June). Insiders' profits, costs of trading, and market efficiency. *Journal of Financial Economics*, *16*(2), 189–212. https://doi.org/10.1016/0304-405X(86)90060-7

Sharpe, W. F. (1964). Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk*. *The Journal of Finance*, *19*(3), 425–442. (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.1964.tb02865.x) https://doi.org/10.1111/j.1540-6261.1964.tb02865.x

Sharpe, W. F. (1966). Mutual Fund Performance. *The Journal of Business*, *39*(1), 119–138. (Publisher: University of Chicago Press) Retrieved from [2025-06-30]https://www.jstor.org/stable/2351741

Shefrin, H. (2002). *Beyond Greed and Fear: Understanding Behavioral Finance and the Psychology of Investing*. Oxford University Press. (Google-Books-ID: hX18tBx3VPsC)

Sherwood, M., & Pollard, J. (2018). *Responsible Investing: An Introduction to Environmental, Social, and Governance Investments* (1st ed.). New York: Routledge.

Sparkes, R. (2002, December). Ethical Investment: Whose Ethics, Which Investment?

Business Ethics: A European Review, *10*, 194–205.

Sustainalytics. (2020, July). *Sustainalytics Makes its ESG Ratings and Data on More Than 4,000 Companies Publicly Accessible on its Website.* Retrieved from [2025-09-07]https://www.sustainalytics.com/esg-news/news-details/2020/07/22/sustainalytics-makes-its-esg-ratings-and-data-on-more-than-4-000-companies-publicly-accessible-on-its-website

Sustainalytics. (2024, May). *ESG Risk Ratings - Methodology Addendum.* Sustainalytics. Retrieved from [2025-09-08]https://www.sustainalytics.com/docs/default-source/meis/definitionsofmeis.pdf?sfvrsn=8e7552c0_9

Sustainalytics. (2025a, September). *ESG Risk Ratings.* Retrieved from [2025-09-08]https://www.sustainalytics.com/corporate-solutions/esg-solutions/esg-risk-ratings

Sustainalytics. (2025b, September). *Top-Rated Companies.* Retrieved from [2025-09-08]https://www.sustainalytics.com/corporate-solutions/esg-solutions/top-rated-companies

Thomson, B., & Bastit, B. (2024, October). Sustainability Brief: TISFD Faces An Uphill Road To Bring Clarity To Social Reporting. *S&P Global*. (Archived: https://web.archive.org/web/20241102152005/https://www.spglobal.com/ratings/en/research/arti sustainability-brief-tisfd-faces-an-uphill-road-to-bring-clarity-to-social-reporting-13294289) Retrieved from [2025-06-20]https://www.spglobal.com/ratings/en/research/articles/241021-sustainability-brief-tisfd-faces-an-uphill-road-to-bring-clarity-to-social-reporting-13294289

Titman, S., Wei, K., & Xie, F. (2004). Capital investments and stock returns. *Journal of Financial and Quantitative Analysis*, *39*(4), 677–700.

UN PRI. (2023, January). *An introduction to responsible investment | PRI.* (https://web.archive.org/web/20250508233009/https://www.unpri.org/introductory-guides-to-responsible-investment/what-is-responsible-investment/4780.article) Retrieved from [2025-06-09]https://www.unpri.org/introductory-guides-to-responsible-investment/what-is-responsible-investment/4780.article

UN PRI. (2025, May). *Reporting Framework glossary.* (https://archive.is/zi5Me) Retrieved from [2025-06-09]https://www.unpri.org/reporting-and-assessment/

reporting-framework-glossary/6937.article

UNPRI. (n.d.). *What is responsible investment? | Introductory guide | PRI*. Retrieved from [2023-11-26]https://www.unpri.org/introductory-guides-to-responsible-investment/what-is-responsible-investment/4780.article

US SIF. (2021). *2020 Report on US Sustainable and Impact Investing Trends* (Tech. Rep. No. 5). (US) Sustainable Investment Forum. Retrieved from [2025-01-16]https://www.ussif.org/files/Trends/2020_Trends_Onepager_Alternatives.pdf

van Duuren, E., Plantinga, A., & Scholtens, B. (2016, October). ESG Integration and the Investment Management Process: Fundamental Investing Reinvented. *Journal of Business Ethics*, *138*(3), 525–533. https://doi.org/10.1007/s10551-015-2610-8

Verheyden, T., Eccles, R. G., & Feiner, A. (2016, June). ESG for All? The Impact of ESG Screening on Return, Risk, and Diversification. *Journal of Applied Corporate Finance*, *28*(2), 47–55. (Number: 2) https://doi.org/10.1111/jacf.12174

von Wallis, M., & Klein, C. (2015, October). Ethical requirement and financial interest: a literature review on socially responsible investing. *Business Research*, *8*(1), 61–98. https://doi.org/10.1007/s40685-014-0015-7

Appendices

Appendix 1. TISFD forum for developing globally consistent disclosure framework

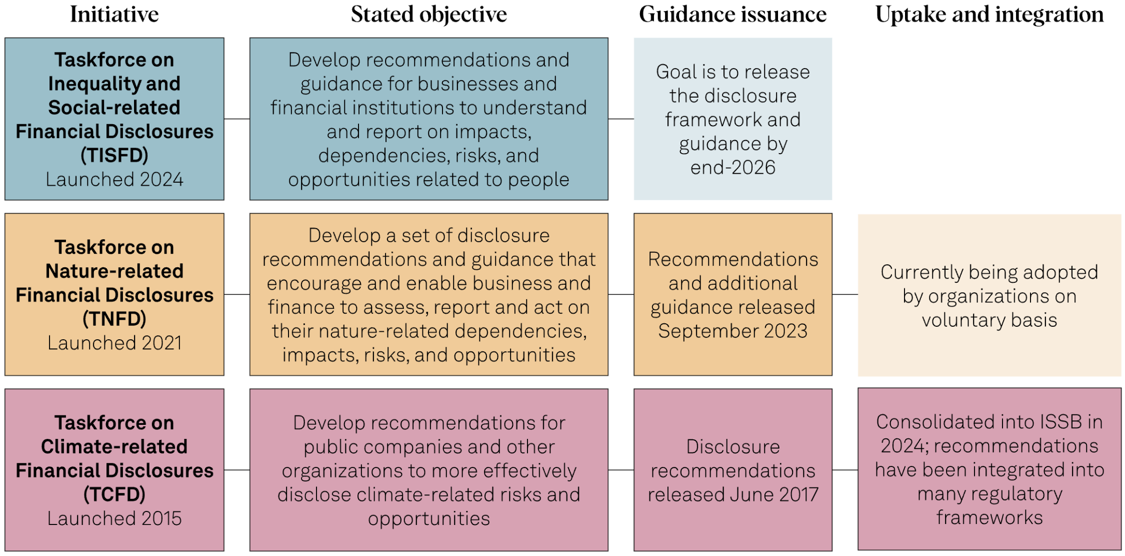

Figure 12. TISFD follows in the footsteps of environmental task forces (Thomson & Bastit, 2024).

Appendix 3. Summary of different ESG investment styles

Table 9. Summary of CSR investment strategies models (Global Sustainable Investment Alliance, 2013).

	CSR/ESG investment strategies
ESG Integration	The systematic and explicit inclusion by investment managers of ESG factors into financial analysis
Corporate engagement & shareholder action	Employing shareholder power to influence corporate behaviour, including through direct corporate engagement (i.e.,communicating with senior management and/or boards of companies), filing or co-filing shareholder proposals, and proxy voting that is guided by comprehensive ESG guidelines.
Norms-based screening	Screening of investments against minimum standards of business or issuer practice based on international norms such as those issued by the UN, ILO, OECD and NGOs (e.g. Transparency International
Negative/exclusionary screening	The exclusion from a fund or portfolio of certain sectors, companies, countries or other issuers based on activities considered not investable Exclusion criteria (based on norms and values) can refer, for example, to product categories (e.g., weapons, tobacco), company practices (e.g., animal testing, violation of human rights, corruption) or controversies
Best-in-class/positive screening	Investment in sectors, companies or projects selected for positive ESG performance relative to industry peers, and that achieve a rating above a defined threshold.
Sustainability themed/thematic investing	Investing in themes or assets specifically contributing to sustainable solutions - environmental and social - (e.g., sustainable agriculture, green buildings, lower carbon tilted portfolio, gender equity, diversity).
Impact investing	Investing to achieve positive, social and environmental impacts - requires measuring and reporting against these impacts, demonstrating the intentionality of investor and underlying asset/investee, and demonstrating the investor contribution.
Community investing	Where capital is specifically directed to traditionally underserved individuals or communities, as well as financing that is provided to businesses with a clear social or environmental purpose. Some community investing is impact investing, but community investing is broader and considers other forms of investing and targeted lending activities.

Appendix 2. Large cap and Geographic Biases in ESG ratings

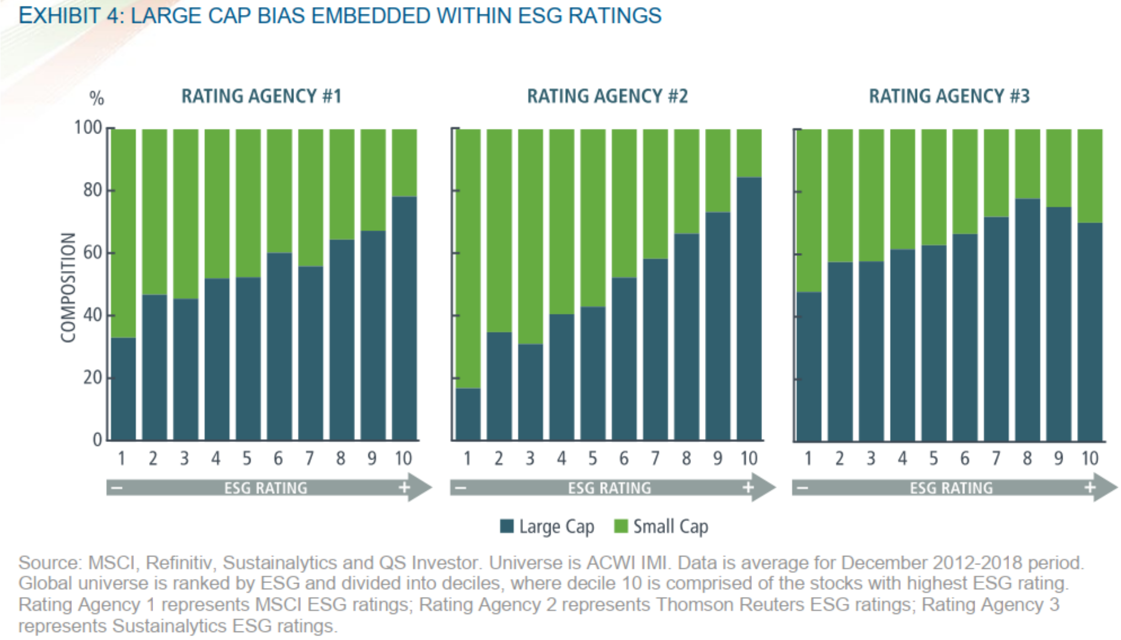

Figure 13. Large cap bias present in ESG ratings(LaBella et al., 2019, 5).

Appendix 3. Summary of different ESG strategies

Type of ESG Strategy	Papers	Results	Stock Universe
ESG Overall and Component Alpha	Deconstructing ESG Ratings Performance (Giese, Nagy and Lee, 2020)	Long/short portfolios from sorting on ESG, E, S, G (and other more specific key issue scores) come with **positive active return** and positive alpha.	MSCI AC World Index IMI (2013-2019).
	Foundations of ESG Investing: How ESG Affects Equity Valuation, Risk, and Performance (Giese, Lee, Melas, Nagy and Nishikawa, 2019)	"the **performance advantage** of higher ESG-rated companies is visible across the entire universe"	Several MSCI universes including Europe and US
ESG Momentum Alpha	Can ESG add alpha? An analysis of ESG Tilt and Momentum Strategies (Nagy, Kassam and Lee 2016)	ESG momentum strategies generate **positive alpha**.	MSCI World Index (2008-2015)
	How Markets Price ESG: Have Changes in ESG Scores Affected Stock Prices? (Giese and Nagy, 2018)	A long/short ESG momentum strategy shows a **positive alpha**.	MSCI World Index (2009-2018), MSCI Emerging market index (2013-2018).
ESG Combined Alpha	ESG for All? The Impact of ESG Screening on Return, Risk and Diversification (Verheyden, Eccles and Feiner, 2016)	Excluding stocks with lowest ESG scores leads to **improved returns**, lower volatility, and lower tail risk.	Large and Mid Cap Global and Developed (2010-2015).

Table 10. Overview of Papers on ESG and Positive Alpha.

Appendix 4. Factor regressions in full

ESG-MOM with H=1

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.034
Model:                            OLS   Adj. R-squared:                 -0.034
Method:                 Least Squares   F-statistic:                     1.283
Date:                Mon, 29 Sep 2025   Prob (F-statistic):              0.281
Time:                        06:51:07   Log-Likelihood:                 248.35
No. Observations:                  77   AIC:                            -484.7
Df Residuals:                      71   BIC:                            -470.6
Df Model:                           5
Covariance Type:                  HAC
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const      -9.798e-05      0.001     -0.116      0.908      -0.002       0.002
Mkt-RF         0.0089      0.032      0.281      0.779      -0.053       0.071
SMB            0.0416      0.051      0.818      0.413      -0.058       0.141
HML            0.0127      0.046      0.276      0.782      -0.077       0.103
RMW            0.0533      0.061      0.877      0.381      -0.066       0.172
CMA           -0.0616      0.070     -0.879      0.379      -0.199       0.076
==============================================================================
Omnibus:                       17.602   Durbin-Watson:                   2.246
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               24.349
Skew:                          -0.968   Prob(JB):                     5.16e-06
Kurtosis:                       4.960   Cond. No.                         78.8
==============================================================================

Notes:
[1] Standard Errors are heteroscedasticity and autocorrelation
robust (HAC) using 6 lags and without small sample correction
```

ESG-MOM with H=3

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.050
Model:                            OLS   Adj. R-squared:                 -0.019
Method:                 Least Squares   F-statistic:                    0.9856
Date:                Mon, 29 Sep 2025   Prob (F-statistic):              0.433
Time:                        06:51:08   Log-Likelihood:                 246.13
No. Observations:                  75   AIC:                            -480.3
Df Residuals:                      69   BIC:                            -466.4
Df Model:                           5
Covariance Type:                  HAC
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const      -9.049e-05      0.001     -0.105      0.916      -0.002       0.002
Mkt-RF         0.0309      0.034      0.905      0.366      -0.036       0.098
SMB            0.0186      0.048      0.385      0.700      -0.076       0.113
HML           -0.0219      0.033     -0.666      0.505      -0.086       0.043
RMW            0.0056      0.059      0.095      0.924      -0.111       0.122
CMA           -0.0484      0.049     -0.989      0.323      -0.144       0.047
==============================================================================
Omnibus:                        0.349   Durbin-Watson:                   1.998
Prob(Omnibus):                  0.840   Jarque-Bera (JB):                0.267
Skew:                           0.142   Prob(JB):                        0.875
Kurtosis:                       2.936   Cond. No.                         78.1
==============================================================================

Notes:
[1] Standard Errors are heteroscedasticity and autocorrelation
robust (HAC) using 6 lags and without small sample correction
```

ESG-MOM with H=6

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.098
Model:                            OLS   Adj. R-squared:                  0.030
Method:                 Least Squares   F-statistic:                     1.488
Date:                Mon, 29 Sep 2025   Prob (F-statistic):              0.206
Time:                        06:51:09   Log-Likelihood:                 224.87
No. Observations:                  72   AIC:                            -437.7
Df Residuals:                      66   BIC:                            -424.1
Df Model:                           5
Covariance Type:                  HAC
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const         -0.0033      0.001     -3.903      0.000      -0.005      -0.002
Mkt-RF         0.0014      0.045      0.031      0.975      -0.086       0.089
SMB           -0.0070      0.037     -0.186      0.852      -0.080       0.066
HML            0.0286      0.039      0.728      0.466      -0.048       0.106
RMW           -0.2015      0.091     -2.208      0.027      -0.380      -0.023
CMA           -0.0183      0.049     -0.375      0.708      -0.114       0.078
==============================================================================
Omnibus:                        7.427   Durbin-Watson:                   1.937
Prob(Omnibus):                  0.024   Jarque-Bera (JB):                8.037
Skew:                          -0.497   Prob(JB):                       0.0180
...

Notes:
[1] Standard Errors are heteroscedasticity and autocorrelation robust (HAC) using 6 lags and without small sample correction
```

ESG-MOM H=12

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.115
Model:                            OLS   Adj. R-squared:                  0.042
Method:                 Least Squares   F-statistic:                     1.644
Date:                Mon, 29 Sep 2025   Prob (F-statistic):              0.162
Time:                        06:51:09   Log-Likelihood:                 207.11
No. Observations:                  66   AIC:                            -402.2
Df Residuals:                      60   BIC:                            -389.1
Df Model:                           5
Covariance Type:                  HAC
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.0003      0.001      0.322      0.748      -0.002       0.002
Mkt-RF        -0.0441      0.036     -1.241      0.215      -0.114       0.026
SMB            0.1332      0.058      2.309      0.021       0.020       0.246
HML           -0.0219      0.043     -0.513      0.608      -0.106       0.062
RMW            0.0244      0.114      0.214      0.830      -0.198       0.247
CMA            0.0950      0.075      1.270      0.204      -0.052       0.241
==============================================================================
Omnibus:                        0.340   Durbin-Watson:                   1.716
Prob(Omnibus):                  0.844   Jarque-Bera (JB):                0.290
Skew:                          -0.154   Prob(JB):                        0.865
...

Notes:
[1] Standard Errors are heteroscedasticity and autocorrelation robust (HAC) using 6 lags and without small sample correction
```

ESG-VOL, W=18

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.018
Model:                            OLS   Adj. R-squared:                 -0.071
Method:                 Least Squares   F-statistic:                    0.5205
Date:                Mon, 29 Sep 2025   Prob (F-statistic):              0.760
Time:                        06:51:10   Log-Likelihood:                 194.10
No. Observations:                  61   AIC:                            -376.2
Df Residuals:                      55   BIC:                            -363.5
Df Model:                           5
Covariance Type:                  HAC
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.0038      0.001      3.752      0.000       0.002       0.006
Mkt-RF        -0.0125      0.013     -0.932      0.351      -0.039       0.014
SMB            0.0370      0.044      0.841      0.400      -0.049       0.123
HML           -0.0217      0.035     -0.621      0.535      -0.090       0.047
RMW           -0.0242      0.051     -0.471      0.638      -0.125       0.077
CMA           -0.0094      0.050     -0.191      0.849      -0.107       0.088
==============================================================================
Omnibus:                       59.938   Durbin-Watson:                   2.176
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              425.464
Skew:                           2.655   Prob(JB):                     4.09e-93
Kurtosis:                      14.799   Cond. No.                         80.3
==============================================================================

Notes:
[1] Standard Errors are heteroscedasticity and autocorrelation robust (HAC) using 6 lags and without small sample correction
```

ESG-VOL with W=24

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.041
Model:                            OLS   Adj. R-squared:                 -0.057
Method:                 Least Squares   F-statistic:                     3.014
Date:                Mon, 29 Sep 2025   Prob (F-statistic):             0.0189
Time:                        06:51:11   Log-Likelihood:                 171.91
No. Observations:                  55   AIC:                            -331.8
Df Residuals:                      49   BIC:                            -319.8
Df Model:                           5
Covariance Type:                  HAC
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.0046      0.001      3.454      0.001       0.002       0.007
Mkt-RF        -0.0142      0.025     -0.572      0.567      -0.063       0.035
SMB            0.0162      0.051      0.319      0.749      -0.083       0.116
HML           -0.0331      0.046     -0.726      0.468      -0.123       0.056
RMW           -0.0608      0.065     -0.939      0.348      -0.188       0.066
CMA           -0.0159      0.059     -0.269      0.788      -0.132       0.100
==============================================================================
Omnibus:                       55.232   Durbin-Watson:                   2.111
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              381.807
Skew:                           2.561   Prob(JB):                     1.24e-83
Kurtosis:                      14.848   Cond. No.                         85.5
==============================================================================

Notes:
[1] Standard Errors are heteroscedasticity and autocorrelation robust (HAC) using 6 lags and without small sample correction
```

ESG-VOL W=30

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.148
Model:                            OLS   Adj. R-squared:                  0.049
Method:                 Least Squares   F-statistic:                     2.519
Date:                Mon, 29 Sep 2025   Prob (F-statistic):             0.0437
Time:                        06:51:12   Log-Likelihood:                 172.84
No. Observations:                  49   AIC:                            -333.7
Df Residuals:                      43   BIC:                            -322.3
Df Model:                           5
Covariance Type:                  HAC
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.0031      0.001      3.256      0.001       0.001       0.005
Mkt-RF        -0.0527      0.027     -1.964      0.049      -0.105      -0.000
SMB           -0.0371      0.029     -1.267      0.205      -0.094       0.020
HML            0.0123      0.022      0.552      0.581      -0.031       0.056
RMW           -0.0572      0.044     -1.311      0.190      -0.143       0.028
CMA           -0.0822      0.060     -1.378      0.168      -0.199       0.035
==============================================================================
Omnibus:                        4.640   Durbin-Watson:                   1.958
Prob(Omnibus):                  0.098   Jarque-Bera (JB):                3.551
Skew:                           0.501   Prob(JB):                        0.169
Kurtosis:                       3.857   Cond. No.                         87.9
==============================================================================

Notes:
[1] Standard Errors are heteroscedasticity and autocorrelation robust (HAC) using 6 lags and without small sample correction
```

ESG-VOL W=36

OLS Regression Results

Dep. Variable:	y	R-squared:	0.023
Model:	OLS	Adj. R-squared:	-0.109
Method:	Least Squares	F-statistic:	0.4899
Date:	Mon, 29 Sep 2025	Prob (F-statistic):	0.782
Time:	06:51:12	Log-Likelihood:	145.78
No. Observations:	43	AIC:	-279.6
Df Residuals:	37	BIC:	-269.0
Df Model:	5		
Covariance Type:	HAC		

	coef	std err	z	P>|z|	[0.025	0.975]
const	0.0012	0.001	0.896	0.370	-0.001	0.004
Mkt-RF	-0.0118	0.019	-0.613	0.540	-0.050	0.026
SMB	-0.0290	0.040	-0.727	0.467	-0.107	0.049
HML	-0.0010	0.027	-0.036	0.971	-0.055	0.053
RMW	-0.0157	0.052	-0.303	0.762	-0.117	0.086
CMA	-0.0280	0.060	-0.467	0.640	-0.146	0.090

Omnibus:	0.387	Durbin-Watson:	1.938
Prob(Omnibus):	0.824	Jarque-Bera (JB):	0.541
Skew:	0.016	Prob(JB):	0.763
Kurtosis:	2.451	Cond. No.	86.8

Notes:
[1] Standard Errors are heteroscedasticity and autocorrelation robust (HAC) using 6 lags and without small sample correction

ESG-VOL with UMD and DEF with W=24

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.035
Model:                            OLS   Adj. R-squared:                 -0.153
Method:                 Least Squares   F-statistic:                    0.5728
Date:                Sat, 27 Sep 2025   Prob (F-statistic):              0.773
Time:                        07:01:19   Log-Likelihood:                 135.56
No. Observations:                  44   AIC:                            -255.1
Df Residuals:                      36   BIC:                            -240.9
Df Model:                           7
Covariance Type:                  HAC
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.0040      0.001      3.222      0.001       0.002       0.006
Mkt-RF        -0.0091      0.015     -0.624      0.533      -0.038       0.020
SMB            0.0718      0.066      1.089      0.276      -0.057       0.201
HML           -0.0237      0.038     -0.619      0.536      -0.099       0.051
RMW           -0.0147      0.082     -0.180      0.857      -0.175       0.145
CMA           -0.0037      0.055     -0.067      0.946      -0.112       0.104
UMD            0.0166      0.031      0.541      0.589      -0.044       0.077
DEF           -0.0274      0.064     -0.431      0.667      -0.152       0.097
==============================================================================
Omnibus:                       44.193   Durbin-Watson:                   2.194
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              195.822
Skew:                           2.426   Prob(JB):                     3.00e-43
Kurtosis:                      12.125   Cond. No.                         88.5
==============================================================================
```

ESG-MOM with H=6

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.162
Model:                            OLS   Adj. R-squared:                 -0.001
Method:                 Least Squares   F-statistic:                     1.421
Date:                Sat, 27 Sep 2025   Prob (F-statistic):              0.227
Time:                        07:01:19   Log-Likelihood:                 135.44
No. Observations:                  44   AIC:                            -254.9
Df Residuals:                      36   BIC:                            -240.6
Df Model:                           7
Covariance Type:                  HAC
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const         -0.0027      0.001     -2.495      0.013      -0.005      -0.001
Mkt-RF         0.0090      0.045      0.197      0.843      -0.080       0.098
SMB           -0.0244      0.070     -0.351      0.726      -0.161       0.112
HML            0.0104      0.033      0.317      0.752      -0.054       0.075
RMW           -0.2607      0.125     -2.092      0.036      -0.505      -0.016
CMA            0.0578      0.094      0.613      0.540      -0.127       0.242
UMD            0.0068      0.046      0.148      0.882      -0.083       0.096
DEF            0.0074      0.032      0.236      0.814      -0.054       0.069
==============================================================================
Omnibus:                        1.267   Durbin-Watson:                   2.015
Prob(Omnibus):                  0.531   Jarque-Bera (JB):                0.480
Skew:                          -0.011   Prob(JB):                        0.787
Kurtosis:                       3.511   Cond. No.                         88.5
==============================================================================

Notes:
[1] Standard Errors are heteroscedasticity and autocorrelation robust (HAC) using 6 lags and without small sample correction
```