Jasmin Rinne

The Performance of ESG Exchange Traded Funds

School of Finance
Master's thesis in Finance
Master's degree Programme in Finance

Vaasa 2023

UNIVERSITY OF VAASA
School of Accounting and Finance
Author: Jasmin Rinne
Title of the Thesis: The Performance of ESG Exchange Traded Funds
Degree: Master of Science in Economics and Business Administration
Subject: Finance
Supervisor: Janne Äijö
Year: 2023 **Pages:** 77

ABSTRACT:

This paper presents an overview of the two growing market trends, passive asset management and Sustainable and responsible investing. Sustainable and responsible investing through exchange-traded funds (ETFs) has grown immensely in the past five years. ESG investing incorporates Environmental, Social, and Governance (ESG) considerations into the investment decision-making process, whereas an exchange-traded fund (ETF) represents a passively managed investment that replicates a specific index or financial asset. The purpose of this thesis is to examine whether combining these two would lead to higher returns and be profitable in the long term in the U.S. market. The sample time covers the period from December 2012 to December 2022.

After the overview, recent literature on the impact of ESG levels on long-term performance is discussed in greater detail. In terms of the literature on combining passive asset management and ESG, there are still relatively few studies and conflicting findings. However, according to most studies, the results are not significant enough to prove the impact of higher ESG performance. An analysis of the differences between portfolio performance and abnormal returns generated by different ESG ratings is conducted. A variety of factor models are used to determine alpha, including CAPMs, Fama-French three-factor models, Carhart four-factors, and Fama-French five-factor models. Moreover, the Sharpe ratio is used to expand the analysis of risk-corrected performance.

Lastly, empirical results reveal that during the study period, each ETF portfolio had positive alphas, suggesting the possibility of abnormal performance, ESG portfolio returns are not significant and yielded less than the market benchmark. The results of this study show that ESG ETFs have become a competitive investment option with the potential to incorporate personal values into investments but lack significant returns.

KEYWORDS: ETF, ESG, ETF, Sustainability, Responsible investing, SRI, Passive asset management

VAASAN YLIOPISTO
School of Laskentatoimen ja rahoituksen yksikkö
Tekijä: Jasmin Rinne
Otsikko: The Performance of ESG Exchange Traded Funds
Koulutus: Master of Science in Economics and Business Administration
Maisteriohjelma: Finance
Työn ohjaaja: Janne Äijö
Valmistumisvuosi: 2023 **Pages:** 77

TIIVISTELMÄ:

Passiivisen omaisuudenhoidon sekä kestävän ja vastuullisen sijoittamisen käyttö on lisääntynyt merkittävästi viimeisten vuosien aikana erityisesti pörssinoteerattujen rahastojen (ETF) avulla. ETF:t ovat passiivisesti hoidettuja pörssilistattuja sijoitusrahastoja, jotka jäljittelevät tiettyä indeksiä tai rahoitusvälinettä. ESG-sijoittaminen sisällyttää ympäristöön, yhteiskuntaan ja hallintotapaan liittyvät kysymykset osaksi sijoitusten päätöksentekoprosessia. Tämän Pro Gradu -tutkielman tarkoituksena on selvittää, tuottaako näiden kahden yhdistäminen epänormaalia tuottoa ja onko se pitkällä aikavälillä kannattavaa Yhdysvaltain markkinoilla. Tutkimusjakso kattaa ajanjakson joulukuusta 2012 joulukuuhun 2022. Tutkielmassa käytetään Morgan Stanley Capital International -kestävyysluokitusta vastuullisuuskertoimien analysoimiseen.

Tutkielmassa käsitellään yksityiskohtaisemmin viimeaikaista kirjallisuutta ESG-luokitusten vaikutuksesta pitkän aikavälin suorituskykyyn. Passiivisen omaisuudenhoidon ja ESG:n yhdistämistä käsittelevästä kirjallisuudesta on vielä suhteellisen vähän tutkimuksia ja ristiriitaisia havaintoja. Useimpien tutkimusten mukaan tulokset eivät kuitenkaan ole riittävän merkittäviä, jotta ne osoittaisivat korkeamman vastuullisuuskertoimen vaikutuksen parempaan suorituskykyyn. Suoritetaan analyysi salkun suoriutumisen ja eri vastuullisuuskertoimien tuottamien poikkeavien tuottojen eroista. Alfan määrittämiseen käytetään erilaisia faktorimalleja, kuten CAPM-malli, Fama-French kolmifaktorimallia, Carhart nelifaktorimallia ja Fama-French viisifaktorimallia. Lisäksi Sharpe-suhdetta käytetään riskikorjatun suorituskyvyn analyysiin, jotta tutkimusta saataisiin laajennettua.

Lopuksi empiiriset tulokset paljastavat, että tutkimusjakson aikana jokainen ETF-salkku tuotti positiivisia alfoja, mikä viittaa epänormaalin tuottojen mahdollisuuteen. ESG-salkun tuotot eivät kuitenkaan ole merkittäviä ja tuottivat vähemmän kuin markkinoiden vertailuarvo. Tämän tutkimuksen tulokset osoittavat, että ESG ETF:istä on tullut kilpailukykyinen sijoitusvaihtoehto, jolla on potentiaalia sisällyttää henkilökohtaisia arvoja sijoituksiin, mutta joiden tuotto ei ole merkittävä.

KEYWORDS: ETF, ESG, ETF, Sustainability, Responsible investing, SRI, Passive asset management

Contents

Figures

Tables

Abbreviations

AP	Authorized Participants
AUM	Assets Under Management
CAPM	Capital Asset Pricing Model
CSR	Corporate Social Responsibility
C4F	Carhart Four-Factor model
ESG	Environmental, Social and Governance
ETF	Exchange Traded Fund
EUROSIF	European Sustainable Investment Forum
FF3F	Fama-French Three-Factor model
FF5F	Fama-French Five-Factor model
ISS	Institutional Shareholder Services
MSCI	Morgan Stanley Capital International
NASDAQ	National Association of Securities Dealers Automated
NAV	Net Asset Value
NYSE	New York Stock Exchange
OLS	The Ordinary Least Square
PRI	Principles for Responsible Investment
SIF	The Forum for Sustainable and Responsible Investment
SPDR	S&P Depositary Receipts Trust Series 1
SRI	Socially Responsible Investing
UN	The United Nations
US	The United States

1 Introduction

Passive asset management and sustainable and responsible investment have grown their popularity among investors in recent years. Given the growth's exponential nature, it intends to continue as a trending and growing topic in the future.

Sustainable and responsible investing considers environmental, social, and governance (ESG) criteria in investment decisions to improve returns and manage risks, while also allowing investors to align their personal values with financial profits. (Auer, 2016; Stevenson & Tuckwell, 2019). The beginning of ESG investing dates back to the 1960s, which made investors more conscious of environmental matters. Sustainable and responsible investing has experienced remarkable growth over the past decade. Assets incorporated in ESG portfolios have grown from $6 trillion to $121 trillion globally in the years 2006-2021. Since ESG funds have become available for everyone, individual investors are more and more interested in investing in them. Yet, most of the products are still actively managed mutual funds, but sustainable passive investing is growing through ETFs and index funds. (Bioy & Lamont, 2018; SSGA, 2020.)

The importance of ESG issues has got more attention since the COVID-19 pandemic and global warming. Investors are now prioritizing issues such as income and gender inequality, employee well-being, and sustainable development, expanding the investment spectrum beyond environmental concerns to social and governance factors. Hence, the overall responsibility has been taken into consideration more broadly, which has encouraged companies to include social and governance factors in their goals and business models to meet the broader responsibility and demands of investors. This change will not only give companies a competitive edge going forward but also highlights how quickly the world is changing. Additionally, the rapid expansion in recent years suggests that investors and investment firms are increasingly focusing on ESG issues. (SSGA, 2020b.)

Furthermore, the market for passive asset management has grown considerably because of the rise in the popularity of ETFs. ETFs are passively managed funds that are replicating the performance of a specific index or financial instrument. These have transformed into a more cost-effective and risk-free investment form than mutual funds. Following the 2008 financial crisis, ETFs have experienced exponential growth and surpassed a remarkable 600% increase. With over 7000 ETFs globally ETFs reached $10 trillion in assets under management as of 2021. (Statista, 2022). Most of the ETFs are benchmarked to a specific index, which allows them to have lower fees and more diversified alternative ways to invest passively. However, it is essential to consider the risks connected with ETFs, such as counterparty risk, trading fees, market risk, closure risk, and potential tracking inaccuracies. (ETF database, 2017; Hill et al., 2015.)

Previously sustainable and responsible investing was expensive and only limited to high-priced mutual funds or required a deep understanding of ESG when selecting securities. The major reason for the increasing demand is that sustainable ETFs are substantially lower in price than sustainable mutual funds as well as offering greater transparency and accuracy for investors. (BlackRock, 2020.) The demand for ESG ETFs has risen by the increased awareness of ESG issues, which have enhanced the availability and quality of ESG data. By the end of 2019, the number of funds globally was over 300 and that number has since increased to almost 1200. Despite the exponential growth, ESG ETFs still represent only a fraction of the total assets invested in ESG and ETFs. However, ESG ETFs have gained significant attraction which is reflecting the overall trend in the U.S. market and that has changed the focus of academic papers on analyzing the results of this combination. (BlackRock, 2020; The Asset, 2022; Trackinsight, 2022.)

The research on combining passive asset management and ESG are still relatively minor, and the results of academic findings have been contradictory over time on whether integrating ESG into investment decisions affects financial performance. Large number of previous research on the performance of ESG have focused on the nexus between ESG and mutual funds. This thesis approaches the subject from the perspective of Exchange

Traded Funds due to the rapid changes in popularity and future possibilities. Although, various studies are debating on the results and many of the recent studies are stating that ESG-screened indexes have been beating their market peers. The question remains whether it has an actual effect and whether could it be profitable in the long term.

The aim of this study is to respond to the gap in the literature by examining the impact of ESG ratings on the performance of Exchange Traded Funds (ETFs), focusing specifically on whether ESG ETFs can demonstrate long-term profitability as an investment option. The creation of a portfolio continues the reader's introduction to sustainable and ESG investing practices and ETFs as an investment instrument. The empirical part of this study examines whether ESG has an impact on the financial performance of ETFs and can ESG ETFs demonstrate long-term profitability as an investment option. Although, it fails to succeed in that. The empirical analysis of this thesis shows that each ETF portfolio were able to display positive alphas, indicating the possibility of abnormal performance. Yet, the ESG portfolio returns are not significant and yielded less than the market benchmark. To support this, previous studies have also found that there is a connection between the level of ESG scores and the returns yet the returns between high ESG scores and low are not significant enough (Halbritter and Dorfleitner, 2015). Overall, although ESG ratings might influence portfolio performance, the results indicate that ESG rating is not the main driving factor of abnormal returns. There are many different possible causes behind the non-significant results which will be discussed in the following chapters.

The overall result of this study shows the way for future research on ESG ETFs. Future ETF research in the context of ESG can be further improved by leveraging growing knowledge and adapting the approach accordingly. Hence, the empirical research should be continued for more accurate results.

1.1 Purpose of the Study

The purpose of the study is to examine the impact of ESG ratings on the performance of Exchange Traded Funds (ETFs), focusing specifically on whether ESG ETFs can demonstrate long-term profitability as an investment option. The study concentrates on the U.S. market, and it covers the period from December 2012 to December 2022. Considering the exponential growth during the past few years and the relatively minor research on combining passive asset management, this thesis responds to a new on-trend topic. As the results of academic findings have been contradictory over time, there can be seen a gap in the literature.

ETFs, being a cost-effective investment option compared to mutual funds, provide an opportunity to achieve higher returns. However, with the increasing popularity of ESG investing, companies are compelled to enhance their practices to meet the growing demand. This situation can potentially result in companies creating misleading information to portray themselves in a more favorable light when it comes to ESG ratings. Moreover, larger corporations have the financial resources to invest in improving the quality and quantity of their ESG disclosures, which can distort the outcomes. Investments that are not subject to restrictions offer greater flexibility and are less exposed to heightened sensitivity. As a result, the outcomes can be contradictory or inconclusive.

Some studies suggest that higher ESG ratings are associated with higher profitability and lower risk incidents, indicating a potential for abnormal returns. However, other studies have found no significant alpha or abnormal returns linked to ESG ratings. Since the evidence is stronger in favor of insignificance, we assume that there is no significant difference in abnormal returns between high and low-ESG-rated ESG ETFs. Hence, the null hypothesis for this study is as follows:

H0: High ESG rating does not lead to statistically significant abnormal returns for passive ESG ETFs compared to ESG ETFs with low ratings.

H1: High ESG rating will lead to statistically significant abnormal returns compared to ESG ETFs with low ratings.

Although we assume that ETFs with high ESG rating will not lead to abnormal returns, it is believed that they will show higher long-term profitability than lower ESG rating ETFs. The changes firms make in the area of environmental, social, and governance, will improve their overall success. Therefore, their strategic thinking and value creation are for the longer term. In addition, the use of ETFs gives a bigger cost advantage over time with a considerably small risk due to their structure and better stakeholder relations. This can lead to lower costs of capital which enables these companies to access finance on more favorable terms, lowering their overall financial load. ESG investing attracts a larger number of investors, which increases the demand for ESG ETFs and could increase their profitability. To be more specific, since we want to establish whether higher ESG ratings are associated with higher profitability in long-term compared to low ESG ratings in ETFs, we can propose the following alternative hypothesis:

H2: ETFs with high ESG ratings show higher profitability in the long term.

H3: ETFs with low ESG performance show lower profitability in the long term.

In conclusion, the assumptions are based on the existing literature. The literature shows contradictory findings concerning the influence of ESG ratings on abnormal returns and long-term profitability. The hypotheses aim to discover and provide insights into the relationship between ESG ratings and the financial performance of ESG ETFs by considering the possibility of abnormal returns and differences in profitability between high and low-ESG-rated ETFs.

1.2 Structure of the Study

The structure of the study is as follows. This thesis will have eight main chapters. The first section introduced the topic and gave a background for the study. It presented why the topic is relevant now and in the future. Consequently, it reasoned how investing in ESG Exchange Traded Funds will progress in the future and why the subject is relevant now.

Chapters two and three will introduce the key terms of the study. The concept of Exchange Traded Fund and Sustainable and Responsible Investing will be presented in greater detail, and it will give a broader understanding of the study and its purpose. Chapter two will present in general, how ETFs work and a brief overview of the history of how they have developed. Also, the risks and costs of the instrument will be discussed. Chapter three gives a general perspective of the various terms concerning sustainable and responsible investing. More closely, the term ESG is introduced and how it performs when using it as a part of an investing strategy. Lastly, the usage of ESG ratings as a part of ESG investing and their issues are presented.

The fourth chapter will discuss the relevant previous literature and research regarding the topic and chapter five will give a description of the data and descriptive statistics. The methodology of this thesis is presented in chapter six. It will introduce four common asset pricing models the CAPM, the Fama-French three, and the five-factor model as well as Carhart four-factor model. Also, the Sharpe Ratio is presented as discussing and reasoning the designated theories.

Lastly, chapter seven will reveal the empirical analysis and results and the last chapter presents a discussion and the conclusions from the main findings.

2 ETF – Exchange Traded Fund

This chapter presents the Exchange Traded Funds (ETFs), the investment instrument of this study, and its discrepancies from mutual funds. In the first subchapter, the risks and costs of ETFs are discussed. Following the history and development of Exchange Traded Funds, and their current status is presented.

An Exchange Traded Fund is a passively managed investment fund, which replicates a specific index or a given financial instrument. ETFs consist of a basket of securities. These securities can be traded on an exchange just as shares of stocks. Therefore, ETF prices change throughout the day and can be sold or bought in a short amount of time, whereas mutual funds allow investors to redeem shares once a day after the market closes. Consequently, purchase and sale transactions will determine the price of ETFs. There are both actively managed and passively managed ETFs yet the majority of the ETFs are passively managed. The passively managed ETFs are not trying to outperform the benchmarked index, which differs from the actively managed ETFs, where the focus is to outperform the index. Nowadays, there is a significant number of different ETFs, which provide a wide selection of alternatives to invest in. They can contain e.g., commodities, currencies, bonds, and equities based on the replicated indexes. These can be divided into international, regional, and country-specific ETFs. In addition, there are specialty asset classes and fixed-income ETFs. (Bodie, Kane & Marcus, 2018; Lydon, 2009; Puttonen & Repo, 2007.)

ETF's structure has features of open- and closed-end funds. As an open-end fund permits the creation and redemption, the closed-end fund permits the shares to be traded on exchanges. However, there is no direct link between ETFs and end investors, and the trading of ETFs occurs through brokers in the secondary market. The brokers either conduct the trade with the market-maker or on exchange. The authorized participants (APs) are the only participants that can create or redeem new ETF shares in the primary market. They often are e.g., a large broker or market-maker authorized to directly transact with the ETF sponsor. The APs exchange the redeemed ETF shares for securities.

This creation and redemption mechanism keeps the deviation of prices from the net asset value (NAV) small. (Ben-David, Franzoni & Moussawi, 2017; ECB, 2018; Hill, Nadig & Hougan, 2015.)

ETFs are divided into two different groups: physical ETFs and synthetic ETFs. The physical ETFs are holding usually all the stocks from the index, with weights, aiming to mimic the return of the benchmark index. In addition, many of these lend portfolio securities to other borrowers. The engaging of securities lending is a way for physical replication to yield additional profit. The Synthetic ETFs use derivative contracts to track the index e.g., total return swaps on the benchmark index. Thus, this allows the synthetic ETF to have higher tracking performance and to replicate more illiquid markets. Equivalently, the risk is higher in synthetic ETFs, and they are also exposed to the risk of default of the counterparty in the derivative contract. This is discussed in the risks and costs chapter in greater detail. Synthetic replication has also gained more popularity in Europe in comparison to the U.S. markets. (ECB, 2018; Ramaswamy, 2011; Ben-David et al., 2017.)

In recent years, investors have demonstrated more interest in ETFs due to the advantages they offer compared to mutual funds. ETFs are traded on an exchange, and it consequently enhances the liquidity and lowers the fees of these instruments. They can be sold short, held on margin, or optioned just like stocks. There is no distribution, account servicing, or management fees included in the expense ratio, due to the structure of ETF. Hence, ETFs benefit from tax advantages. The low turnover and redemption mechanism are enabling factors for investors to minimize the tax implications when exchanging ETFs. The assets can be traded without creating capital gain taxes because investors can sell or exchange their shares with other traders, without selling the shares from the underlying portfolio. (Abner, 2016; Bodie et al., 2018; Hill et al., 2015; Lettau & Madhavan, 2018.)

In contrast with all the underlying benefits, there are also problems with ETFs. Especially, when the markets are volatile, measuring the NAV of the ETF portfolios tracking less

liquid assets may be difficult. Therefore, the number of ETFs that have less dealers, has a risk of prices fluctuating notably. This becomes relevant if the dealers flee from the market during a period of uncertainty. The prices can deviate from NAV for short periods because investors trade ETFs as securities. Consequently, these deviations induce a diminishing in the offered cost advantage in ETFs. (Bodie et al., 2018.)

2.1 The history and development of ETFs

The history of ETFs dates back to 1990 when the Toronto 35 Index was introduced to the Toronto Stock Exchange. Although, the more known date is in January 1993, when the first US-listed ETF called a spider, made its debut on the American Stock Exchange. The SPDR, the so-called spider is benchmarked to the S&P 500 Index. (Meziani, 2016.)

However, according to Meziani (2016), the ETF industry did not spark significant interest amongst investors until 2005-2006. Puttonen and Repo (2007) note that in 2004-2005 e.g., the rise of oil share- and emerging markets, assisted the ETFs breakthrough. By the end of 2005 assets under management (AUM) had grown to 301$ billion and were doubled in the following two years (Meziani, 2016). Additionally, the financial crisis of 2008-2009 brought more interest towards index products, and by the end of 2009, almost a thousand ETF products were trading on the U.S. exchanges (Lydon, 2010; Puttonen & Repo, 2007).

The growth of ETFs has been enormous. According to Ben-David, Franzoni, and Moussawi (2017) in 2016 over 10% of the market value of securities traded on the U.S. Stock Exchange and approximately 30% of overall daily trading volume in the U.S., were represented by ETFs. Today there are over 7000 ETFs globally. In addition, global assets under management held by ETFs have reached $10 trillion in 2021. (SSGA, 2020; Statista, 2022.)

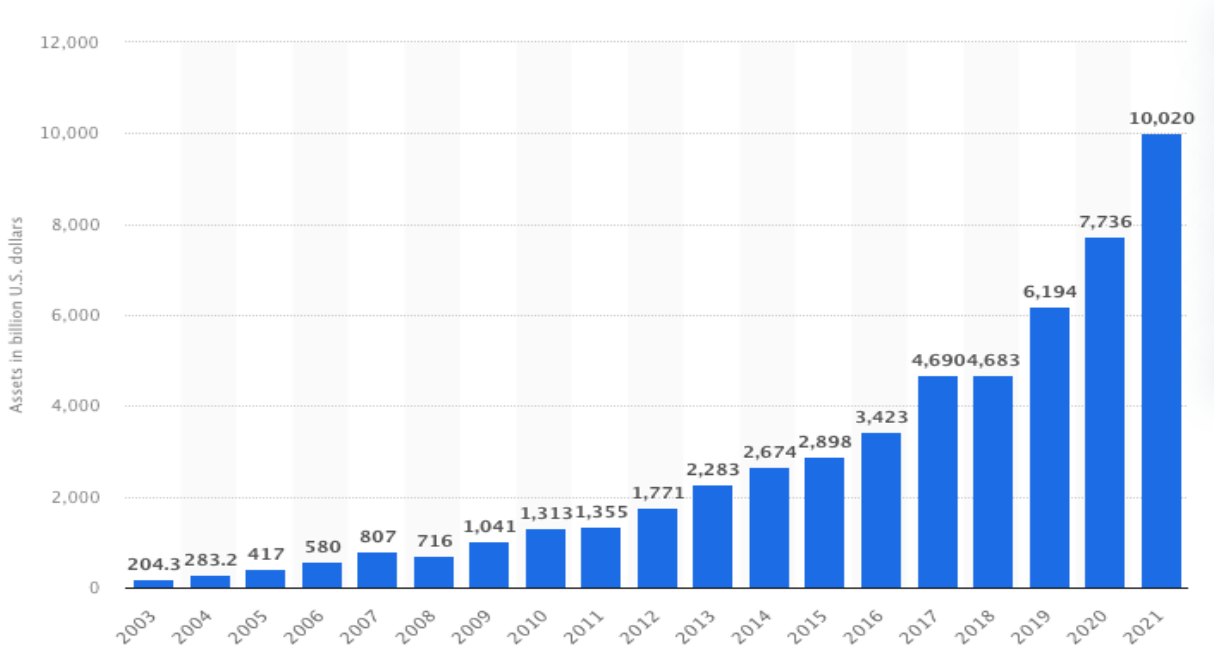


Figure 1. Growth of Global ETFs over time

Source: Statista (2022).

Figure 1. demonstrates how the assets under management have developed globally over the years from 2003 to 2021. The growth has been exponential after the 2008 financial crisis. For the last decade, the growth is over 600% and after 2015 it is nearly over 250%. The total assets in 2020 amounted to over $7.7 trillion, which means the growth over the past years has been over $2.5 trillion (Statista, 2022).

2.2 Risks and costs of ETFs

As mentioned above, ETFs have a cost advantage over mutual funds and are accounted for as a less risky asset class. However, ETFs do have both transparent and hidden fees, and are not immune to the risks. These fees reduce the total return of the investments. The riskiness is defined by diversification, and the fees depend on the structural differences of the asset class. (ETF database, 2017; Fidelity, 2021.) These risks and costs should be considered when examining the returns.

Market risk is one of the risks and costs that affect all asset classes. As mentioned above, ETFs replicate a specific index or a given financial instrument. The performance of these underlying assets and the market itself determines the performance of the ETF. Hence, this means that market risk is impossible to be directly mitigated. Nevertheless, investors can seek different ways to allocate the assets in a way that indirectly mitigates the riskiness. (ETF database, 2017.) Also, another overall risk, the closure risk, should be considered when managing ETFs. The risk of the fund closing, is higher with ETFs due to the ratio between the low level of investors in relation to a large number of ETFs. Therefore, that might lead to failure of attracting those investors. This means that the AUM should be considered when evaluating the riskiness of the fund. The environment of ETFs is constantly changing as well as its regulations. The change in regulation can force the fund to close e.g., due to position limits. (Hill et al., 2015.)

Although ETFs benefit from cost advantage due to their structure. Yet the ownership of the ETF portfolio creates cots to the investor, which is also defined as a trading risk. ETFs are traded on an exchange and therefore include commission fees. This can be an issue if the assets deposited into an ETF are small and frequent. Hence, the expense ratio will be higher, and the investment will lose its cost-effectiveness. These and the other expenses such as capital gain and interest income taxes, creation and redemption fees should be considered into the total trading costs when mitigating the risks and overall costs. (ETF database, 2017; Fidelity, 2021.)

The essential factor for investors is the possibility to redeem the assets. ETFs liquidity can be observed in two ways in this matter. As an average daily trading volume which also provides the historical trading frequency, and as a gauge of the potential amount to be traded in ETFs which is called implied liquidity. Since ETFs are not in direct interaction with end-investors, the implied liquidity screens the trading conditions more accurately. In general, liquidity is not an issue for the largest and most popular ETFs and consequently, investors tend to have expectations of high ETF liquidity in all market situations. This is likely to occasion pressure for investors to sell if liquidity becomes

weaker due to the popularity based on the liquidity advantage that ETFs have. Extreme volatility is one of the market situations where market participants may have to pay more to provide liquidity. (ECB, 2018; ETF database, 2017.)

The counterparty risk arises in ETFs using synthetic replication or exploiting securities lending. Synthetic ETF exploit total return swaps to track the benchmark index to gain exposure as mentioned earlier. The risk, in this case, can be reduced by collateralizing the swap exposure of the fund. As a response to the higher risk, synthetic ETFs have lower tracking error and fees than their physical peers. Equivalently, physical ETFs can yield more profit by using securities lending to borrowers. In both cases, the investor might encounter the risk of loss if the swap counterparty or the borrower defaults on its commitments. In the case of a market downturn, market situations become more stressful, and the counterparty risk increases. This means that larger redemptions are made which can induce pressure for ETF to sell collateral securities and downward pressure on asset prices. (ECB, 2018; ETF database, 2017.)

Usually, ETFs tracking the same index or sector may not perform similarly due to the divergent allocation in their underlying securities. When replicating the target index there is a possibility that the weights and securities differ from the benchmarked index. This creates the composition risk that is considerably higher when tracking a specific sector due to the variety of companies and segments in the industry. The tracking error is similar to this. It becomes accurate because the arrangement fees, tax management, and timing of dividends leads to the ETF not performing as the target index. The physical ETFs encounter higher tracking errors than synthetic replications. (ECB, 2018; ETF database, 2017.)

Lastly, the popularity among ETFs has grown massively and hence increased the number of new ETFs launching. This can create the hype risk which can be caused by a herd mentality of market participants or in this case the bull market. The upward trend of markets creates a bubble that feeds investors to mimic other market participants when

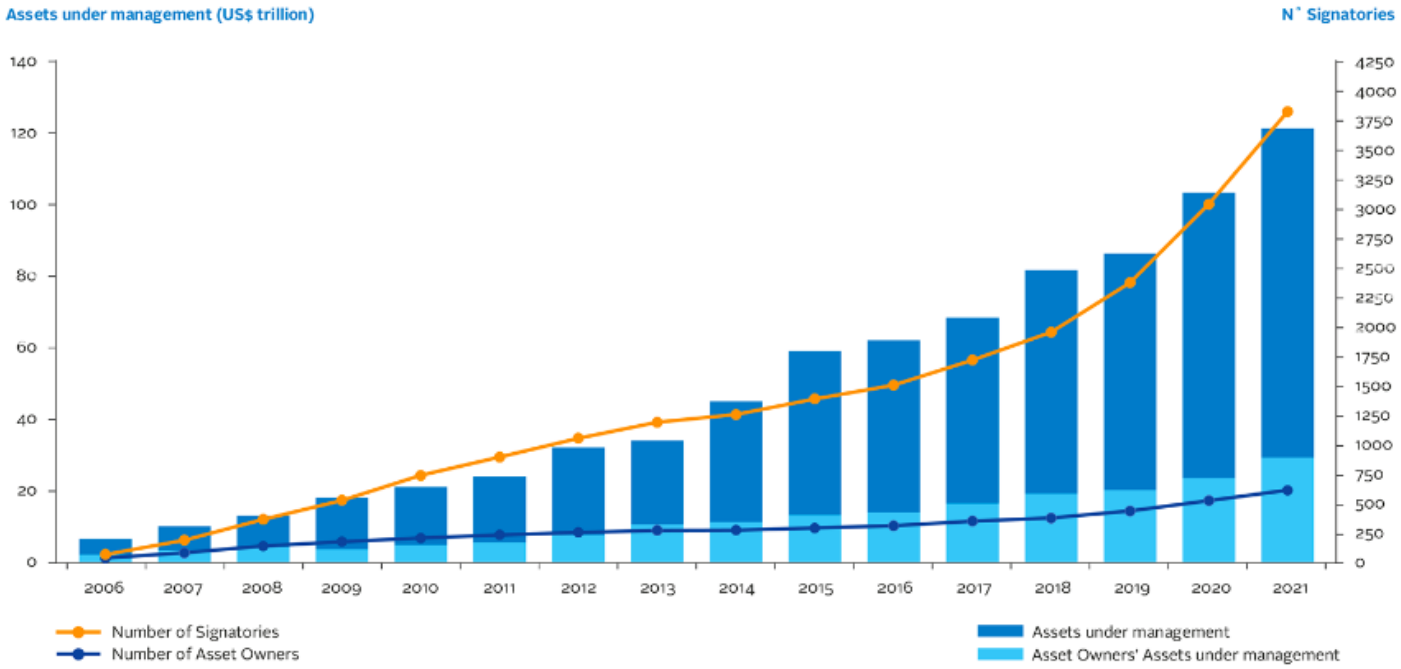
seeking the best profits. Hence, investors might end up investing in less profitable investments as they easily forget to examine the context of the target investment. (ETF database, 2017.) Hill et al. (2015) note that misunderstanding of how the ETF works is also the biggest risk when investing in them. The risk is minimized when market participants educate themselves and have a clear investment strategy that is followed despite the surrounded herd behavior (ETF database, 2017).

3 Sustainable and Responsible Investing

In this chapter, the concept of sustainable and responsible investing is presented. The following subchapters focus on introducing the environmental, social, and governance criteria more broadly. Consequently, explaining ESG investing, its ratings, and issues as well as presenting the MSCI ESG Rating.

The terminology over sustainability and responsibility lack consensus and most of them slightly overlap each other. Sustainable and responsible investing utilizes both traditional investment approaches together with environmental, social, and governance (ESG) factors to improve the return and risk profile. Moreover, it seeks competitive risk-adjusted returns in long term. BlackRock (2020) defines sustainable investing as an umbrella for investment strategies and ESG is the data or gauge of information for sustainable investment solution. Studies have examined if non-financial performance has a significant impact on financial performance revealing the potential for generating superior returns and identifying investment risks. Additionally, sustainable investing has changed its perspective from the focus on restricting investments and divestiture to highlight the positive impact on risk assessment and return potential of companies. (Arabesque, 2020.)

Sustainable and responsible investing can be divided into specific categories. These contain e.g., Socially Responsible Investing (SRI), Impact investing and ESG investing. Each category executes a different strategy. Socially Responsible Investing is based on ethical and moral criteria and aims to avoid negatively affecting industries or stocks e.g., alcohol, gambling, tobacco, and weapons. (Arabesque, 2020; Kanuri, 2020.) Whereas Impact investing is seeking the financial return by investing in companies or projects with a goal of social and/or environmental change. The third is ESG investing which integrates ESG factors with financial factors in the investment decision-making process. It will be introduced in more detail in the following subchapters. (Caplan, Griswold & Jarvis, 2013; MSCI, 2020b.)



Source: UN PRI (2021).

Figure 2. PRI Growth Globally 2006-2021.

Figure 2. above demonstrates how the awareness of ESG issues has grown exponentially. Assets under management have grown from $6 trillion to $121 trillion in the years 2006-2021. In the past five years, the growth has doubled. The global pandemic over the past few years have accelerated growth. The amount of PRI investor signatories reached 26% growth only in the year 2021. When comparing the development between the years 2018-2019 and 2019-2020, there is a notable difference from approx. 5% to 25% growth. The growth in 2021 solely was again almost 20%. (SSGA, 2020; UN PRI, 2021.)

3.1 ESG – Environmental, social, and governance

Environmental, Social, and Governance criteria are defined as a measurement tool to indicate the performance of non-financial factors when evaluating corporate behavior. ESG criteria are used for improving the risk- and return profile. (Stevenson & Tuckwell, 2019.) ESG disclosure provides the investors with a possibility to incorporate personal

values into the financial profits (Auer, 2016). The beginning of ESG investing dates back to the 1960s, when investors started to exclude different industries and stocks from their portfolios, incorporated with e.g., tobacco or apartheid regime (MSCI, 2020b).

ESG concentrates on three subareas regarding environmental, social, and governance aspects in companies' functions. The environmental aspect reflects environmental consciousness that appears through the company's involvement in waste management, energy efficiency, and biotechnology. Social issues incorporate human rights, consumer privacy, employment diversity, and data security. Lastly, the governance factor comprises issues in board structure, financial reporting, corruption, and business ethics, and culture, etc. Figure 3. presents examples of these common ESG issues. (Auer, 2016; HSBC, 2019.) Despite the same data, the data providers choose to assess the information differently by considering different dimensions in each factor. Therefore, it can be questioned whether the ratings are comparable with the ESG indices. (Dorfleitner, Halbritter & Nguyen, 2015.)

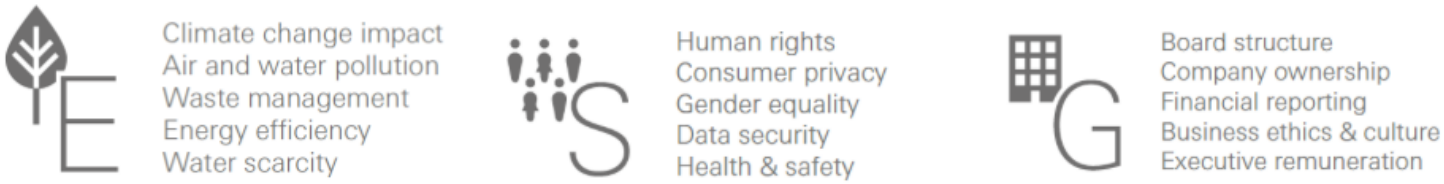


Figure 3. Examples of common ESG issues
Source: HSBC (2019).

To meet the investors growing demand, a large number of institutions e.g., the Principles for Responsible Investment (PRI), an UN-supported initiative offer a set of principles that assist the incorporation of ESG factors into ownership and investment decisions. It aims to enhance the return and risk management with wider profits of society. There are six principles for investors. These principles encourage investors to e.g., include ESG issues into different ownership practices, investment analysis, and decision-making process. In

addition, when cooperating with others, promoting and reporting their actions, the effectiveness of incorporating ESG issues will intensify. (UN PRI, 2020.)

Another large institution, the Forum of Sustainable and Responsible Investment (SIF), has classified different approaches to integrate ESG into their investment processes. Those approaches are divided into seven categories mentioned below: Negative screening, ESG integration, Corporate engagement, Norms-based screening, Positive/best-in-class screening, Sustainability-themed investing, and Impact investing. Negative screening is the oldest and most used strategy where the concentration is in excluding companies, sectors, and countries based on their negative influence e.g., tobacco, animal testing, nuclear energy. However, ESG integration and corporate engagement have increased their popularity with more active management by investors. Figure 4. presents the definition of each strategy in greater detail. The strategies work both individually and in a combination. (Eurosif, 2018; HSBC, 2019.)

Example ESG strategies	Definition
Negative/exclusionary screening	Focused on excluding the "sin" stocks, e.g. cluster munitions, alcohol, tobacco, and gambling
ESG integration	The systematic inclusion by managers of ESG factors into traditional financial analysis
Corporate engagement	The use of shareholder power to influence corporate behaviour
Norms-based screening	Excluding investments that do not meet minimum standards of business practice based on international norms
Positive screening	Including stocks of companies that show positive ESG performance
Sustainability-themed investing	Including stocks of companies that pursue sustainable development
Impact investing	Targeting specific environmental or social outcomes along with financial returns

Figure 4. ESG investment strategies
Source: HSBC (2019).

The importance of ESG issues has got attention since the COVID-19 pandemic, global warming, etc. Issues such as income and gender inequality, the well-being of employees, and sustainable development, are now top goals for investors. This has changed the focus from the ESG's ''E'', that has already been adopted and considered as a trending factor in the investment spectrum, to ''S'' and ''G''. Hence, the overall responsibility is taken into consideration more broadly, which encourages companies to focus more on

their non-financial factors. These will be creating new investment opportunities for the future as the world is changing rapidly. Also, the PRI's strong growth over the past years has shown that the investors and investment companies are turning to focus more on ESG issues. (SSGA, 2020b.)

3.2 Performance of ESG investing

The broad attention that ESG investing has received, has provoked questions about the long-term impact and the ability to outperform other forms of investment. The academic research has increased interest in examining the possibilities of ESG investing. However, the literature regarding the performance of ESG investing is contradictory. This incurs from the divergent results that different ESG strategies and a broad range of ESG scores induce. Also, there are various types of gauges how the impact on performance can be measured e.g., Fama-French multi factors, Sharpe Ratio, etc., which will be introduced in the fourth chapter. (Bauer, Koedijk & Otten, 2005; Derwall, Koedijk & Ter Horst, 2011; MSCI, 2020a.)

The various methods in the construction of the variables explain why there is no distinct consensus between ESG factors and performance (MSCI, 2020a). Equivalently, according to Hill, Ainscough, Shank and Manullang (2007), only a small number of investors have a conclusion of what is considered to be socially responsible, which complicates the measurement of the performance. The broad array of themes concerning each environmental, social, and governance factor diverse so widely that it is improbable, they will have a similar influence on a company's performance. Also, the extent that ESG factors are reflected in their overall ESG scores, or individual E, S, and G scores is debatable. ESG scores are used to evaluate companies based on their environmental, social, and governance practices and to measure how they handle those issues. (MSCI, 2020a.)

Both Sharfman and Fernando (2008) and Clark, Feiner and Viehs (2015) agree that companies with high ESG performance have a lower cost of capital and leverage. Therefore, these companies usually have higher debt relative to their peers. Buallay (2019) states that it is due to the better reputation and social policy of the companies. The companies with prominent environmental practices tend to have less volatile stock performance (Sharfman & Fernando, 2008). Also, Fernando (2008) debates that risk management is one of the indicators in ESG score so, companies with higher ESG scores have lower business risks. Whereas Kumar (2019) supports that high ESG ratings do tend to result in higher profitability, but the alpha remains insignificant and close to zero.

The academic findings give varying results and there are multiple ways to implement ESG issues into investment decisions. However, ESG scores have become the most employed way to indicate the sustainability of companies or investment instruments.

3.2.1 ESG -ratings

The majority of companies both international and domestic have been rated by different data vendors based on environmental, social, and governance performance. ESG ratings are significant information to corporations and investors integrating ESG issues into their investment selection processes. (Dorfleitner et al., 2015; Huber & Comstock, 2017.) Especially, as Crilly et al. (2012) notes that corporations are increasingly being pressured by stakeholders to integrate social responsibility. These rating methodologies are used to constructing ESG scores. The investment instruments' such as ETFs' ESG score is consisted of the company-level competence to handle ESG issues and is the aggregate of these company-level ESG scores. A large number of vendors are providing ESG data such as Bloomberg, MSCI, ISS, Thompson Reuters, Sustainalytics, and Morningstar. (MSCI, 2020b.)

Table 1: ESG Report/Rating Summary Table

Source: Huber & Comstock (2017).

ESG Report Provider	Background	Rating Scale	Methodology	Usage and Reputation
Bloomberg ESG Data	Collects ESG Data for over 9,000 companies Integrated into Bloomberg Equities and Intelligence Services	Out of 100 Provides scores from third-party rating agencies	Looks at 120 ESG indicators	In 2016, Bloomberg had over 12,200 ESG Customers
ISS	ISS's solutions also include climate change data and analytics from its recent acquisition of Climate Neutral Investments ISS QualityScore provides corporate governance reports on over 5,600 public companies	ISS QualityScore: 1-10 Climetrics Score: 1 to 5 green leaves	ISS QualityScore: Covers board structure, compensation/remuneration, shareholder rights, and audit & risk oversight Updated on an ongoing basis	A leading provider
MSCI ESG	Provides ratings for over 6,000 companies and 350,000 equity and fixed income securities	AAA to CCC	Looks at 35 Key ESG Issues Data collected from publicly available sources Companies monitored on an ongoing basis	iShares MSCI EAFE ESG Select ETF and MSCI EM ESG Select ETF Institutional investors, including Legal and General Investment Management, Morgan Stanley, Northern Trust Asset Management, and PIMCO
Sustainalytics	2008 consolidation of DSR, Scoris and AIS Covers over 6,500 companies across 42 sector	Out of 100 Sector/industry based comparison	Looks at industry-specific ESG indicators, covers at least 70 indicators in each industry And, systems to manage ESG risks and disclosure of ESG issues and performance	Strategic relationships with BNY Mellon, City of London Investment Management (CLIM), Columbia Threadneedle, Norwegian Government Pension Fund, and Prudential Fixed Income

Thomson Reuters ESG Research Data	Thomson Reuters acquired Asset4 in 2009 Provides ESG data on over 6,000 companies	Percentile rank scores (available on both percentages and letter grades from D- to A+)	Covers 400 different ESG metrics, electing 178 of the most relevant data points Categories are weighted	Comprehensive database ESG Scores are available on Thomson Reuters Eikon platform

3.2.1.1 The MSCI ESG Rating

MSCI ESG Fund Ratings are designed to assess how well mutual funds and exchange-traded funds (ETFs) withstand economically significant E, S, and G risks. MSCI ESG Fund Ratings are created to rate or screen mutual funds and ETFs from the leader, average, to laggard by measuring the Environmental, Social, and Governance attributes of a fund's underlying assets. They provide more than 600 measures that may be used to assess funds on ESG issues, exposure to climate change, values-based concerns, sustainable impact themes, and metrics for the regulatory EU Sustainable Finance landscape in order to facilitate customization. The indicators are drawn from information that is accessible to the general public, including corporate reports, third-party sources and regulatory filings. (MSCI, 2023.)

The scale is between AAA to CCC. Leader (AAA, AA) has a business that excels at addressing the biggest ESG risks and opportunities in its sector. The average (A, BBB, BB) has a business that, in comparison to its counterparts in the industry, has a patchy or average track record of handling the biggest ESG risks and opportunities. The laggard (B, CCC) has a business that lags behind its competitors in its sector due to high exposure to and poor management of key ESG risks. (MSCI, 2023.)

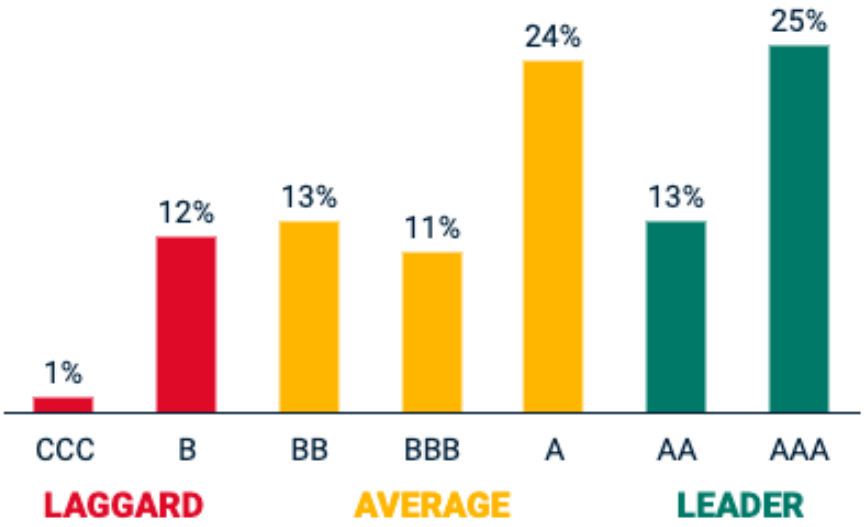


Figure 5. ESG Rating distribution of fund holdings

Source: MSCI (2023).

MSCI Rating methodology looks at 35 Key ESG Issues by evaluating hundreds of data points across 35 ESG Key Issues that concentrate on the junction between a company's core business and the concerns unique to its industry that may present the organization with substantial risks and possibilities. The impact and time frame of the risk or opportunity is considered while weighing the Key Issues. Corporate Governance and Corporate Behavior are evaluated for every company. To determine the Key Issues for each organization and to assign weights for the Environmental and Social Pillars, MSCI's materiality mapping framework is employed. Industry-specific environmental and social key issues are based on environmental or social externalities that could result in sudden costs for a certain industry. (MSCI, 2023.)

The MSCI ESG Rating is essential for evaluating and comparing the sustainability performance of businesses. It has become a useful tool for investors and businesses due to its defined approach and widespread use. The MSCI ESG Rating has a great framework

for assessing and encouraging sustainable company practices across industries as the significance of ESG continues to rise. (MSCI, 2023.)

3.2.2 Problems with ESG-ratings

As previously mentioned, the data and measures regarding ESG rating vary widely between data vendors. SABS (2019) notes that the rapidly growing ESG information available from data vendors and companies differs in quality, quantity, and relevance. One of the issues is discrepancies in regulations and inconsistent jurisdiction between countries that induce inconsistency in the data observed by vendors. Also, various ways to measure and report e.g., employee health or data security, allows companies to provide selective disclosure. Consequently, the same company can get significantly different scores from different vendors. (BlackRock, 2020; Kotsantonis & Serafeim, 2019.)

This leads to the issue with voluntary disclosure. The voluntary disclosure theory argues that companies with high sustainability performance disclose more non-financial information due to the potential increase in market value (Clarkson et al., 2008). According to Christensen et al. (2019) study, it found a nexus between ESG disclosure quantity and variation in ESG ratings. Companies providing more sustainable disclosure, usually have more disagreement in the ESG ratings. When the company's disclosure is voluntary and is not based on a certain regulation, it can be expected. In the line with this Dorfleitner et al. (2015) notes that large companies have a tendency to have high ESG scores due to the increased reporting functions. Greenwashing is one of the issues created by voluntary disclosure as well. Companies aim to highlight their sustainable actions for a better image and therefore lead to embellishing the truth. Due to the misleading information and incorrect product labeling, the reliability of the ESG rating is questioned. (Bloomberg, 2020.)

Konstatonis and Serafeim (2019) consider four inconsistencies in the ESG data available: Incoherence in data, Benchmark options, Imputation of data, and increasing public

information. The incoherence in data is induced by the lack of regulation, a discrepancy in the metrics used by vendors, and disclosure reported by companies. Due to these reasons, it is challenging to compare companies by the ratings given. There are variations in the benchmark options, how companies define the range of good and bad performance. It is usually done by evaluating the peer groups. The lack of transparency regarding their components causes inconsistency and unreliability in the metrics. The third inconsistency considered is the imputation of ESG data, where lack of disclosure is a key issue. These data gaps lead to different imputation approaches by the vendors, which is the primary reason for divergent results in ESG ratings. The data vendors can score these data gaps by using statistical models, giving comparable scores reflecting their peers, or score it as zero. Lastly, publicly available information is increasing significantly. Due to the information expansion, also the disagreement of ESG ratings has expanded. (Dimson et al., 2020; Konstatonis & Serafeim, 2019.)

3.3 ESG Exchange Traded Funds

Despite the inconsistencies mentioned above, investing in passive ESG funds has continued to grow. Especially, increasing demand for ESG index/ETF products has exploded. Previously sustainable investing was expensive and only possible through high-priced mutual funds, or they needed a deep understanding of ESG when picking the securities. Hence, the major reason for the demand is that sustainable ETFs are much lower in price than sustainable mutual funds. These provide more transparency and accuracy for the investors. In addition, the awareness of ESG issues and the development of regulations across industries has improved of ESG data and made it more available to investors. (BlackRock, 2020.)

The first ESG ETF, MSCI USA ESG Select ETF, by iShares was launched in 2002 and by the end of 2009, there were 35 sustainable ETFs globally. Yet, the growth expanded in the following decade, and at the end of 2019, there were over 300, and now around 1200 ESG ETFs. The AUM invested in ESG ETFs, and index funds globally increased to $390

billion at the end of March 2022, meaning that the amount has doubled in 2 years. The past three to five years have been a turning point for sustainable ETFs and index funds. In 2017 the number of new ETF launched outnumbered index funds for the first time. Despite the significant growth, ESG ETFs cover only a fraction of the $121 trillion in assets invested in ESG investing, and $10 trillion in ETFs mentioned earlier. In the U.S. market, ESG ETFs have become a prominent trend and the demand for the product reflects the growth. (BlackRock, 2020; The Asset, 2022; Trackinsight, 2022.)

The pandemic raised the attention to sustainable investing options even more and since then sustainable ETFs have become more competitive option for investing. Companies' and global systems' vulnerability were noticed since companies weren't prepared for shock as global pandemic and how it affected the whole world which highlighted the significance of sustainability and resilience in investment strategies. That resulted investors to be more aware of the fact that it is becoming increasingly clearer that businesses with strong sustainability practices are better equipped to handle crises and reduce risks. Nowadays investors are looking for responsible ETFs as a way to reduce risk and improve portfolio resilience as a result. (Nasdaq, 2021.)

4 Previous literature

With the ongoing growth of ESG investing, the majority of academic papers are aiming to explore the impact of integrating ESG factors into investment decisions on financial performance. However, the research on combining passive asset management and ESG is still relatively minor, and only a few studies have made research on financial performance and value creation of ESG ETFs. Nevertheless, the increasing demand for such investment products has put more focus on analyzing the results of this combination. As mentioned above the results of academic findings have been contradictory over time since the evolving nature of these products. As a result, all the research papers have either short or limited time series of data, or the sample size of the data has been relatively small. This chapter concentrates on examining the performance based on the hypotheses and previous research.

Although most studies have focused on the relationship between ESG and mutual funds, the use of passive investment instruments such as ETFs and index funds has grown significantly. These instruments have changed to more cost-effective investment options than mutual funds. The study conducted by Kumar (2019) highlights how academics and investment institutions are incorporating ESG into portfolio construction to attain abnormal returns. Thus, it is interesting to examine whether it has an actual effect and if it becomes profitable in the long term. In addition, the majority of the ESG ETFs are benchmarked to the ESG index, which means that the properties of the index apply to EGS ETFs as regards examining the risk-return or factor analysis. This allows us to use the results of the performance results of ESG indexes as a part of our study. (Kumar 2019.)

Various studies are debating on the results and many of the recent studies are stating that ESG-screened indexes have been beating their market peers. One of the studies made by Giese, Lee, Melas, Nagy, and Nishikawa (2019) by using MSCI ESG Ratings data and indices as a performance measure, found that a higher ESG rating leads to a lower frequency of idiosyncratic risk incidents indicating higher profitability since their ability to manage significant business risks. Kumar’s (2019) research is supporting this since

there was a positive correlation between the profitability factor and ESG indexes. This suggests that when the ESG score is higher the returns might have a tendency to increase.

Equivalently, Hsu, Liu, Shen, Viswanathan, and Zhao (2018) suggest that investing in an index that includes only firms with high ESG scores, generates higher returns as well as higher social value in comparison to usual ESG indexes. This shows that firms with a high cost of capital would have higher expected returns. Derwall, Guenster, Bauer, and Koedijk (2005) examine a sample of U.S. companies from 1997 to 2003 and the relationship between ecological responsibility and company returns. By employing a high-low strategy and utilizing the Carhart (1997) four-factor model, they find that high-rated firms significantly outperform low-rated firms.

Nevertheless, the reason that led to better results is debatable whether it is an effect of ESG, selection bias, or whether firms with better financial support are investing to enhance their ESG profile. This might be the cause of higher ESG scores and better results. (Giese et al., 2019; Kumar, 2019.) Therefore, large firms that are thriving, have the resources to invest in their outer view through both the quality and quantity of ESG disclosure. This means that the level of the rating might not be the cause of better performance and probably the actual environmental, social, and governance activity is quite low. Manescu (2011) notes that the returns of lower scores could have a similar added premium as sin stocks. Buallay's (2019) study debates this subject and found that the results are inconsistent. There is no valid conclusion on if the disclosure has a negative or positive impact on ESG performance. This highlights the problem with the lack of sufficiently significant evidence.

The research by Halbritter and Dorfleitner (2015) presents a critical analysis of the empirical data concerning the relationship between ESG ratings and financial performance. It emphasizes the necessity of considering the dynamic nature of the SRI market and various ESG rating providers to fully comprehend the connection between financial performance and social performance. They show that there is a nexus between ESG scores and returns. However, the returns between high ESG scores and low are not

significant enough either in overall scores or individual pillars. This is confirmed by Kumar's (2019) study where he is examining the alpha with different asset pricing models from CAPM to Fama-French five-factor model. The results indicate no significant alpha with any of the factor analyses. Hsu et al. (2018) also suggest similar results with top and bottom ESG scores, and Mill (2006) finds no evidence of either outperformance or underperformance. He acknowledges that the ESG integration induces no risk-adjusted returns. In addition, other studies by Statman and Glushkov (2009) and Kempf and Osthoff (2007) find positive abnormal returns but not statistically significant enough.

Contrary to previous papers, Kanuri (2020) shows that ESG ETFs underperformed the U.S. market in terms of both absolute and risk-adjusted performance measures. The paper examines the performance of ESG ETFs in comparison to US and global equity markets. The study period was from 2005 to 2019 and the benchmark portfolios representing the US and international equity markets consistently outperformed the ESG portfolios throughout the period, even though ESG portfolios had both periods of outperformance and times of underperformance. The paper by Limkriangkrai, Koh, and Durand (2017) states that the level of ESG ratings has no significant effect on risk-adjusted returns. They created a high ESG score portfolio, low ESG score portfolio, and portfolios focusing only on the individual E, S, and G components, where each portfolio had at least two different industries. The results indicate no abnormal returns on portfolios based on ESG.

By taking a different approach, Derwall et al. (2011) highlights the complicated and dynamic nature of socially responsible investment (SRI) and its effects on asset prices and investment performance. They found evidence that stocks of companies with high environmental and social scores outperform those with low scores, but there is also evidence that over time, the possibilities diminish. In addition, they have not consistently found higher returns on environmentally conscious portfolios, highlighting the need for additional research. Since this performance can be seen only momentary, the results might be interrelated to a specific time frame, or the strategy used. (Dimson et al., 2020.)

Also, a paper by Filbeck, Holzhauer, and Zhao (2014), used the same type of methodology as Derwall et al. (2011) and added active extension strategies where e.g., stocks with low ESG scores are shorted. They found similar results that enhanced the potential abnormal returns.

Based on various studies, the performance of ESG ETFs has changed over time. It is partially related to the development progress in that period. Not until after the 2008 financial crisis, the growth expanded in ESG investing. Moreover, most of the ESG ETFs have been launched in the past five years. This means that these have not experienced a substantial bear market until 2020. (Hale, 2020.) Bebchuk, Cohen, and Wang (2013) note that ESG portfolios yielded positive abnormal returns in the years 1990-1999. Yet, these turn insignificant in the following years 2000-2008. The abnormal returns were possible due to the mispricing of the ESG stocks but as time evolved market participants learned gradually more about themselves and the functioning of ESG factors.

When in the 2000s, the attention was turned more closely to governance issues. Borgers, Derwall, Koedijk, and Ter Horst (2013) find similar results where indexes based on ESG data performed higher abnormal returns from 1992 to 2004 while those returns turned insignificant after that till 2009. After 2009, Kanuri (2020) detected that ESG ETFs performed similarly to the global index but slightly underperformed the U.S. index. Hence, the performance has been dependent on the period of time and the change that has taken place. There is a large amount of evidence that ESG ETFs have bet the market index over specific period of time and especially during the bull market.

Dimson et al. (2020) note that as a long-term investment, there is a lack of sufficiently significant evidence that ESG integration enhances the returns or lowers risk based on their ratings. Yet, they also find no signs of underperformance. The same results were found by Sarkar, Datta, Mukherjee, and Hannigan's (2016) study, where ESG ETFs slightly underperformed but gave quite similar results compared to SPY in the long-term horizon. They examine the performance between the SPY and various ESG ETFs. After two years,

the difference grows bigger as the horizon expands. This can be noticed e.g., from the cumulative returns from 2005 to 2019. For ESG ETFs it was around 30-60% depending on the strategy employed and for IWT it was 240% and DGT 90%. Equivalently, the returns in the two-year horizon were almost the same, and already in the four-year horizon, the spread was 45% between ESG ETFs and SPY. (Kanuri, 2020; Meziani, 2016.)

However, Sarkar et al. (2016) suggests that ESG ETFs had been yielding abnormal returns per unit of risk when we look at the average Sharpe ratio in the longer term. Lopatta and Kaspereit (2014) found that long-term profitability can be increased by sustainable and responsible corporate actions e.g., efficient production and reduction of costs. These in combination with lower legal and compliance expenses will improve the reputation of the firm as well. They also detected that better stakeholder relations will lower the risks which leads high ESG-score firms to lower the cost of capital and enable higher profitability. This is supported by Buallay's (2019) study where he states that this is an outcome of better reputation and social policy.

A gap is clear in the literature about the advantages of ESG investing in long-term using ETFs considering these academic findings being contradictory over and the lack of a thorough long-term study that analyzes the potential for value creation through ESG Exchange Traded Funds.

5 Data

In this chapter, the data of the study is presented. The chapter starts with a data description. This chapter presents the data of empirical testing of this thesis and the description of ESG data, including the data collection process. Lastly, this is followed by descriptive statistics.

5.1 Data description

The data for this thesis is collected from the Thomson Reuters Eikon database (Refinitiv Datastream), whereas the explanatory variables for the regressions were taken from the Kenneth R. French database (Nofsinger and Varma, 2014). The sustainability classes of the ETFs are also determined using the MSCI ESG rating. This research exclusively focuses on U.S. equity ETFs, since the U.S. has the most established ETF market in the entire world having the majority of publicly accessible ETFs, ensuring the data is accurate and sufficiently comprehensive. Additionally, by concentrating on U.S. stock markets, we eliminate any uncertainty in identifying the data for risk characteristics that are frequently recognized to have an impact on stock prices. (Fama & French, 1993, 2015). The sample has been restricted to domestic equity funds with at least 3 years of data to preserve more robust data and covers the period from December 2012 to December 2022.

As Nofsinger and Varma (2014), the monthly closing prices for the equity ETFs are searched from Refinitiv Datastream as a start. Since there is no sustainability data provided by Refinitiv Datastream, the ESG ratings for each ETF were manually obtained during March 2023. All equity ETFs in the sample must obtain their monthly closing prices from Refinitiv Datastream and ESG data from MSCI for all the ETFs. Benchmarks will be identified to determine whether High ESG ETFs outperform Low ESG ETFs.

Biased performance results can emerge from an improper benchmark selection, as highlighted by Cremers et al. (2012). A benchmark most used by the ESG ETFs and with similar risk exposure to the ESG ETF portfolio of this research and to assess the added value is the S&P500 index. Four different S&P500 equity ETFs will form the benchmark market index, S&P 500 ETF portfolio, for the ESG ETF portfolios. To preserve robustness, the S&P500 index is used as an alternative benchmark market index. (Dumitrescu, Järvinen & Zakriya, 2023.)

5.2 ESG data

First, the sustainability levels for the portfolios are defined based on the MSCI ESG rating assessment. The Portfolios will be divided into high and low portfolios like Halbritter and Dorfleitner (2015) use to investigate the link between ESG and financial performance. MSCI ESG Ratings are designed to assess how well exchange-traded funds (ETFs) withstand economically significant ESG risks. The ESG issue affecting the rating could have a considerable impact on the financial worth of the company and, consequently, on the risk-return profile of the company as an investment. The ESG issues are different depending on the industry and the business. The rating has a scale of AAA-CCC. The AAA rating is considered the best and CCC the worst. The scale has been divided into the leaders (AAA, AA), the average (A, BBB, BB), and the laggards (B, CCC). MSCI provides more than 600 measures that may be used to assess funds on ESG issues. (MSCI, 2023.)

Since almost half of the fund holdings are defined as leaders (AAA, AA) in the MSCI fund ESG rating, portfolio A will include ETFs with an above-average ESG rating. ETFs with AAA or AA ratings are considered to have a low ESG risk and will be included in portfolio A (high). Based on the fund holdings in the MSCI ESG rating, ETFs in the average group (A, BBB, BB) held most of the other half's holdings, the portfolio B (low) will be mainly constructed from these. ETFs with A, BB, or BBB and B, CCC ratings are considered to have a poor or average track record of handling the biggest ESG risks and opportunities. (MSCI, 2023.)

5.3 Descriptive statistics

In this sub-chapter, descriptive statistics are presented regarding age, average returns, skewness, kurtosis, and standard deviation (volatility). The descriptive is annualized for presentation purposes.

Table 2 is a summary of the descriptive statistics derived from the collected and organized data. This is followed by a review of the statistical findings and the initial analyses. In addition, the standard deviation and average returns are measured based on their market capitalization (market cap-weighted). The summary of descriptive statistics is introduced as follows.

Table 2: Table of Descriptive statistics

Descriptive statistic

	Nr.	Avg. Age (y)	Avg. Return (%)	Skew	Kurt	Std (%)
S&P 500 ETF Portfolio	4	22	11,72 %	-0,83	0,18	15,87 %
Portfolio A (high)	235	9,65	5,90 %	-0,69	-1,31	11,82 %
Portfolio B (low)	244	10,71	5,38 %	-1,18	0,67	13,81 %

Table 2 will present S&P 500 ETF Portfolio, Portfolio A (High ESG score), and Portfolio B (Low ESG score) with the number of ETFs in the portfolio, the average age, and the return

of the portfolio. Skewness, kurtosis, and standard deviation (volatility) is also presented. The highest average returns can be observed on the S&P 500 ETF Portfolio while portfolios A and B have similar returns, portfolio A being a little above. During the sample period, all the portfolios had positive returns on average. The lowest Std is observed on portfolio A and the highest on the S&P 500 ETF Portfolio. The ages of the portfolios are measured based on the ETF data from Refinitiv DataStream.

The multivariate skew and kurtosis measures are determined. Distribution skewness quantifies the disparity between the symmetrical structure of a normal distribution and the actual distribution of observed values. Negative skewness indicates that the distribution has a longer or fatter tail on the left side, whereas positive skewness suggests a longer or fatter tail on the right side. On the other hand, kurtosis measures the thickness of the tails in a distribution. As Table 2 shows, all the portfolios are skewed to the left and the kurtosis deviates between -1,18-0,67.

6 Methodology

In this chapter, the theoretical framework of the study is presented. The chapter starts by introducing the Capital Asset Pricing Model (CAPM), which is used to evaluate the expected return for assets. This is followed by Fama-French Three, Five Factor models, and Carhart Four Factor model to understand the impact of risk factors. These three models are improvements to each other. Lastly, the Sharpe ratio is presented to indicate the risk-return ratio. The literature uses these asset pricing models for sustainable and responsible-based funds (Bauer et al., 2005; Derwall et al., 2011).

6.1 Capital Asset Pricing Model

The Capital Asset pricing model (CAPM) is a sequence of assumptions regard to nexus between expected return and systematic risk. It is a ground for all asset pricing models. The model was introduced by William Sharpe (1964), John Litner (1965), and Jan Mossin (1966), and build upon Harry Markowitz's (1952) theory on modern portfolio management. The Capital Asset Pricing Model is based on equilibrium, where investors make the decisions so that the portfolio's variance is minimized considering the expected return, and considering variance, the expected return is maximized. A higher expected return is possible by adding additional risk. By summarizing the kernel of the model, the expected return of the asset will exceed the risk-free return by the extend of risk premium, which is calculated by multiplying market premium with the beta of the asset. (Fama & French, 2004; Sharpe, 1964.) The CAPM can be written as follows:

$$E(r_i) = r_f + \beta_i[E(r_M) - r_f] \quad (1)$$

Where, $E(r_i)$ = Expected return of asset i

r_f = Risk-free rate of return

β_i = Beta of the asset i

$E(r_M)$ = Expected return of market portfolio

Source: Sharpe (1964).

The sequence of assumptions is divided into two: investors' behavior and market structure. Investors' behavior assumes that all investors are rational and concentrate on optimizing the mean and variance in their portfolio return. Thus, the optimal risk-return portfolio is the same as the market portfolio. They have a similar scope of timing in investment and all the information available is public, which leads to homogeneous assumptions. The market has no transaction costs or taxes. Furthermore, all capital goods are available, assets are traded in public exchanges, and can be borrowed or lend at a risk-free rate. (Bodie et al., 2018.) These simplified assumptions induce incoherence in the CAPM's empirical results (Fama & French, 2004).

6.2 Fama-French 3-Factor Model

The CAPM fails to explain multiple patterns discovered by researchers in average stock returns. To respond to these failures, Fama and French (1993) developed a formula of the Fama-French 3-Factor Model. The two patterns without explanation were the relationship between average return and size, and between price ratio such as book-to-market ratio, which is referred as the company's value factor. These two factors: size and value explain the portfolio's abnormal returns. Additionally, it can be used to measure portfolio performance and evaluating the cost of capital. (Fama & French, 1993.) The Fama-French 3-Factor Model can be written as follows:

$$R_{it} - R_{Ft} = \alpha_i + \beta_{i,M}(R_{Mt} - R_{ft}) + \beta_{i,s}SMB_t + \beta_{i,h}HML_t + e_{it} \quad (2)$$

Where, $R_{it} - R_{Ft}$ = Portfolio expected excess return

α_i = Securities risk premium

$(R_{Mt} - R_{ft})$ = Excess return on the market portfolio

$\beta_{i,M;s;h}$= Factor coefficients

SMB_t (small minus big) = size factor

HML_t (high minus low) = value factor

e_{it}= Abnormal return

Source: Fama & French (1993).

Fama and French (1995) found evidence that book-to-market equity and the angular coefficient on HML represent a proxy for relative distress. The weaker companies with constant low earnings usually have a higher book-to-market ratio and positive angular on value factor. Whereas the stronger companies with constant higher earnings usually have a low book-to-market ratio and a negative angular coefficient on value factor, which means lower returns on equity. (Fama & French, 1996.) Chan and Chen (1991) present similar evidence on the existing correlation in returns related to relative distress. Hence, the model can predict abnormal returns for the portfolio.

6.3 Fama-French 5-Factor Model

In the 2015 study, Fama and French introduced 5-Factor Model as an improvement to their 3-Factor Model. This was due to the criticism that Fama and French encountered by e.g., Novy-Marx (2013) and Titman, Wei and Xie (2004). They argue that there is incoherence in the model due to the lack of variation in average returns when considering profitability and investment. The Five-Factor Model takes two new factors into the consideration: profitability factor and investment factor. However, after inserting these new factors into the model, they found that the results were better than earlier. (Fama & French, 2015). The Fama-French 5-Factor Model can be written as follows:

$$R_{it} - R_{Ft} = \alpha_i + \beta_{i,M}(R_{Mt} - R_{ft}) + \beta_{i,s}SMB_t + \beta_{i,h}HML_t + \beta_{i,r}RMW_t + \beta_{i,c}CMA_t + e_{it} \quad (3)$$

Where, RMW_t (robust minus week) = profitability factor

CMA_t (conservative minus aggressive) = investment factor

Source: Fama & French (2015).

In the equation above, the RMW factor is constructed from the returns of diversified portfolios including stocks with robust and weak profitability, whereas the CMA factor compares stocks with high investment firms termed as aggressive and low investment firms termed as conservative. Equivalently, RMW and CMA are construed similarly to HML as means of RMW and CMA for small and big stocks.

The main issue with the 5-factor model is that it fails to explain the low average returns on small caps which profit functions are the same as low profitability companies with considerably high investments. Moreover, it has been found that the value factor will lose its purpose for explaining the average returns after these two additional factors are added. (Fama & French, 2015).

6.4 Carhart 4-Factor Model

The Carhart 4-Factor model was an improvement to the Fama-French 3-Factor model by Carhart (1997) since both the 3-Factor model and CAPM fails to explain the continuation of average returns. Jegadeesh and Titman (1993) highlight the one-year momentum anomaly as a limitation in the 3-Factor model since the findings suggest that past twelve-month returns will continue into the next month's meaning that lower returns follow lower returns and higher returns for higher. The Carhart 4-Factor Model can be written as follows:

$$R_{it} - R_{Ft} = \alpha_i + \beta_{i,M}(R_{Mt} - R_{ft}) + \beta_{i,s}SMB_t + \beta_{i,h}HML_t + \beta_{i,r}MOM_t + e_{it}$$

Where, MOM_t (momentum) = momentum factor

Source: Carhart (1997).

In the equation above, the MOM factor represents the difference in excess returns between portfolios of winning and losing stocks over the previous year. The factor assists to ensure more accurate results due to the fund's sensitivity to momentum. Equivalently, the 4-Factor model is construed similarly, to the 3-Factor model. (Carhart, 1997.)

Fama and French (1996) acknowledge that while their three-factor model eliminates the majority of the average return anomalies that are not considered by the CAPM, the persistence of short-term returns is an exception.

6.5 Sharpe ratio

After the creation of CAPM William F. Sharpe continued his work by introducing the reward-to-variability ratio, commonly termed as the Sharpe ratio. The Sharpe ratio is a risk-adjusted performance metric, which is used to evaluate the risk of expected excess return. The risk of the portfolio is measured by the standard deviation of excess return and the reward is equal to the risk premium. (Bodie et al., 2018; Sharpe, 1994 ;1975) Sharpe ratio can be written as follows:

$$Sharpe\ ratio = \frac{R_p - R_f}{\sigma_p} \quad (4)$$

Where, R_p = Return of the asset p

R_f = Risk-free rate of return

σ_p= Standard deviation of the excess return over the risk-free rate of return

7 Empirical analysis and results

The goal of the research is to determine the portfolio's risk-adjusted return. The portfolios in the research are built using ESG scores, and portfolios are compared to determine whether the level of an ESG score has an impact on performance. Four distinct models were used to create a regression model for the portfolios. The results of the capital asset pricing model are presented first. The findings of multi-factor models, including the Fama-French three-factor and five-factor models as well as the Carhart four-factor model, are presented secondly. The discussion and conclusions are presented in the next chapter.

7.1 The results of CAPM

First, the Capital Asset Pricing Model was used to compute the alphas for portfolio A with a high ESG rating, and portfolio B with a low ESG rating. To further enhance comparability and investigate whether there is a substantial difference between the performance, the alpha was also determined for the benchmark ETF portfolio, S&P 500 ETF portfolio.

Table 3: The results of CAPM

Portfolio	Alpha	Rm-Rf	R-square
S&P 500 ETF Portfolio	9,271	0,967***	0,987
	(0,381)	(26,033)	
Portfolio A (high)	3,207	0,670***	0,897
	(-1,726)	(8,358)	
Portfolio B (low)	3,055***	0,804***	0,946
	(-3,357)	(11,782)	

Table 3 above presents the results from the CAPM model for the 2012/12-2022/12 period. The estimates from the OLS regression for each portfolio. All the alphas are shown in percentages and annualized for clarity of interpretation. The significance levels 10%, 5%, and 1% are marked next to numbers as stars * (10%), ** (5%), ***(1%). T-values are shown in the brackets for each coefficient and the R-square represents the goodness of fit of the regression model.

The results as shown in Table 3, provide the risk and performance characteristics of the portfolios. As the results show all the portfolios generated positive alphas but only the low portfolio B had a statistically significant positive alpha of 3.055 ($t = -3.357$, $p < 0.01$), demonstrating an ability to generate excess returns relative to the market. Although Portfolio A generated a higher alpha than Portfolio B did, the lack of statistical significance for alpha of the Portfolio A, suggests that the observed difference between Portfolio A and the market may not be significant. This could occur by chance or other factors that were not considered in the model. Therefore, for these portfolios, the null hypothesis is rejected. In addition, the data indicates that the benchmark group of the S&P 500 ETF portfolio is significantly outperforming the high and low portfolios and producing far greater returns throughout the course of the entire sample period.

Moreover, all portfolios' positive market risk premiums (Rm-Rf) imply a favorable correlation between market returns. The beta coefficients for each portfolio are less than 1.0 means that these are theoretically less volatile than the market. High portfolio A has a significantly lower beta than the low portfolio or the benchmark portfolio indicating lower sensitivity to market movements. Including this stock in a portfolio makes it less risky than the same portfolio without the stock. The R-squared values from 0,897 to 0,987 show that the CAPM model explains a significant amount of the portfolio performance.

7.2 The results from the 3-factor model

In this chapter, results from the Fama- French 3-Factor Model are presented. By implementing the risk factors SMB for size and HML for value in addition to the market risk component, the Fama-French 3-factor model is computed to further Capital Asset Pricing Model. Equivalently to the CAPM, the factor loadings and alphas for portfolio A, portfolio B, and the S&P 500 ETF portfolio were determined.

Table 4: The results from the 3-factor model

Portfolio	Alpha	Rm-Rf	SMB	HML	R-square
S&P 500 ETF Portfolio	9,394	0,984	-0,048**	-0,032	0,964
	(-0,464)	(28,531)	(-2,718)	(-1,452)	
Portfolio A (high)	3,195	0,664***	-0,007	0,005	0,898
	(-1,475)	(6,449)	(-0,129)	(0,071)	
Portfolio B (low)	3,069***	0,822***	-0,061	-0,047	0,964
	(-3,898)	(11,448)	(-1,720)	(-1,061)	

Table 4 above presents the results from the 3-factor model for the 2012/12-2022/12 period. The estimates from the OLS regression for each portfolio. All the alphas are shown in percentages and annualized for clarity of interpretation. The significance levels 10%, 5%, and 1% are marked next to numbers as stars * (10%), ** (5%), ***(1%). T-values are shown in the brackets for each coefficient and the R-square represents the goodness of fit of the regression model.

As a comparison to the CAPM results, the results demonstrate similarly that although all portfolios have positive alphas, only the alpha for low Portfolio B is statistically significant. This shows that portfolio B has consistently outperformed the market, which cannot be explained by only the three factors in the model. However, all portfolios show positive exposure to the market risk premium (Rm-Rf) similar to CAPM. Portfolio B (low) exhibits negative exposure to both SMB and HML factors, in contrast to Portfolio A (high), which has minimal exposure to both. Consequently, portfolio A is less influenced by these factors. S&P 500 ETF portfolio indicates that the portfolio is skewed towards small-cap ETFs since it's the only portfolio with statistical significance at 5% and negative exposure to the SMB factor. The R-squared values between 0,898-0,964 demonstrate that the 3-factor model significantly explains the performance of the portfolio.

7.3 The results from the 4-factor model

In this chapter, results from the Carhart 4-Factor Model are presented. By implementing the risk factor MOM for momentum in addition to the market, size (SMB), and value (HML) risk factors, the Carhart 4-factor model is computed to further the Fama-French 3-factor model. Equivalently to the 3-factor model, the factor loadings and alphas for portfolio A, portfolio B, and the S&P 500 ETF portfolio were determined.

Table 5: The results from the 4-factor model

Portfolio	Alpha	Rm-Rf	SMB	HML	MOM	R-square
S&P 500 ETF Portfolio	9,493	0,998***	-0,046**	-0,041*	-0,020*	0,997
	(-0,490)	(25,630)	(-3,423)	(-2,337)	(-2,256)	
Portfolio A (high)	3,194	0,664***	-0,007	0,005	0,001	0,898
	(-1,353)	(4,717)	(-0,118)	(0,066)	(0,014)	
Portfolio B (low)	3,090***	0,842***	-0,057*	-0,062	-0,042*	0,983
	(-2,667)	(7,565)	(-2,104)	(-1,833)	(-2,367)	

Table 5 above presents the results from the 4-factor model for the 2012/12-2022/12 period. The estimates from the OLS regression for each portfolio. All the alphas are shown in percentages and annualized for clarity of interpretation. The significance levels 10%, 5%, and 1% are marked next to numbers as stars * (10%), ** (5%), ***(1%). T-values are shown in the brackets for each coefficient and the R-square represents the goodness of fit of the regression model.

As the CAPM results, results show that only the low Portfolio B's alpha is statistically significant, even though other portfolios have positive alphas. This indicates the potential for outperformance beyond risk factors while portfolio B alone has consistently outperformed the market, which cannot be explained by the model's four factors. However, all portfolios show positive exposure to the market risk premium (Rm-Rf) similarly as with CAPM and the 3-factor model, implying that the portfolios will usually profit from positive market movements.

Portfolio B (low) exhibits negative exposure to both SMB and HML factors, in contrast to Portfolio A (high), which has minimal exposure to both. Consequently, portfolio A is less influenced by these factors. S&P 500 ETF portfolio and portfolio B indicate that the portfolios are skewed towards small-cap ETFs since the portfolios have statistical significance at the 5% and 10% and negative exposure to the SMB factor. With the MOM factor, certain coefficients being statistically significant and others not, the results show that the portfolios have different exposures to this momentum factor. This shows that recent price patterns and market momentum may have an impact on the performance of the portfolios. The R-squared values between 0,898-0,997 demonstrate that the 4-factor model significantly explains the performance of the portfolio.

7.4 The results from the 5-factor model

This chapter presents the results of the 5-factor model analysis, which aims to explain the excess returns of the portfolios based on five risk factors: the market risk premium (Rm-Rf), the Small Minus Big (SMB) factor, the High Minus Low (HML) factor, the Robust Minus Weak (RMW) factor, and the Conservative Minus Aggressive (CMA) factor. The Fama-French 5-factor model is computed to further the Fama-French 3-factor model. Equivalently to the 3-factor model, the factor loadings and alphas for portfolio A, portfolio B, and the S&P 500 ETF portfolio were determined.

Table 6: The results from the 5-factor model

Portfolio	Alpha	Rm-Rf	SMB	HML	RMW	CMA	R-square
S&P 500 ETF Portfolio	9,517	1,001***	-0,036	0,003	-0,015	-0,057	0,997
	(-0,490)	(25,630)	(-1,661)	(0,145)	(-0,570)	(-1,998)	
Portfolio A (high)	3,255	0,706***	-0,039	-0,004	-0,048	0,040	0,907
	(-1,353)	(4,717)	(-0,481)	(-0,042)	(-0,488)	(0,363)	
Portfolio B (low)	3,073*	0,826***	-0,063	-0,032	-0,006	-0,008	0,964
	(-2,667)	(7,565)	(-1,068)	(-0,505)	(-0,082)	(-0,097)	

Table 6 above presents the results from the 5-factor model for the 2012/12-2022/12 period. The estimates from the OLS regression for each portfolio. All the alphas are shown in percentages and annualized for clarity of interpretation. The significance levels 10%, 5%, and 1% are marked next to numbers as stars * (10%), ** (5%), ***(1%). T-values are shown in the brackets for each coefficient and the R-square represents the goodness of fit of the regression model.

As a comparison to previous results, they similarly demonstrate that although all portfolios have positive alphas only the alpha for low Portfolio B is statistically significant.

The alpha for Portfolio B is no longer statistically significant at the 1% level but it remains significant at the 10% level. This shift in significance shows that the addition of the RMW and CMA factors has affected the link between the alpha and the factors, leading to a statistically weaker association. However, all portfolios show positive and significant exposure to the market risk premium (Rm-Rf) like previous models, implying that the portfolios will usually profit from positive market movements.

The portfolios show different exposure to the five factors, with minimal exposure to SMB and RMW for all portfolios. Portfolio B (low) is negatively impacted by the HML factor, whereas Portfolio A (high) is positively impacted by the CMA factor. The S&P 500 ETF portfolio shows minimal exposure to all the SMB, HML, RMW, and CMA factors, with coefficients close to zero. These results show how the portfolios' risk and return characteristics compare to the benchmark, showing the potential to produce abnormal performance beyond the factors considered. The R-squared values between 0,907-0,997 demonstrate that the 5-factor model significantly explains the performance of the portfolio while there might be more features or factors not captured in the model that will affect these returns.

7.5 Analysis of Risk-Adjusted Performance

To conduct a more thorough analysis of the abnormal performance of each portfolio, the Sharpe measures are examined to analyze the alphas already stated. All the alphas are shown in percentages and annualized for clarity of interpretation. Furthermore, the risk-adjusted performance increases as the value increases. Table 7 shows the alphas for the Sharpe ratio, Capital asset pricing model (CAPM), Fama-French three-factor (FF3F) and five-factor (FF5F) models as well as the Carhart four-factor model (C4F). All the alphas are shown in percentages and annualized for clarity of interpretation.

Table 7: Portfolio performance measures

	Sharpe Ratio	CAPM	FF3F	C4F	FF5F
S&P 500 ETF Portfolio	0,44	9,27	9,39	9,49	9,52
Portfolio A (high)	0,13	3,21	3,20	3,19	3,26
Portfolio B (low)	0,08	3,06	3,07	3,09	3,07

Table 7 above shows that the S&P 500 ETF Portfolio has the highest risk-adjusted performance, as shown by the higher Sharpe Ratio. In terms of risk-adjusted performance, both the high and low portfolios have underperformed the S&P 500 ETF Portfolio due to lower Sharpe Ratios.

8 Discussion and conclusions

Given the exponential nature of the growth, it is expected that the trend and growth of passive asset management as well as sustainable and responsible investment will continue in the future. A comprehensive analysis of these two significant financial market trends has been done in this thesis. The hypotheses aim to provide insights into the relationship between ESG ratings and the financial performance of ESG ETFs. They are considering the possibility of abnormal returns, and differences in profitability between high and low-ESG-rated ETFs. The theoretical background is presented to explain the divergent outcomes and the recent literature is utilized as data when estimating the hypotheses proposed in this study.

A majority of academic papers have explored the impact of integrating ESG factors into investment decisions on financial performance, with a focus on mutual funds rather than passive asset management instruments like ETFs. However, the growing popularity of passive investment options has increased the demand for analyzing the performance of ESG ETFs. ETFs, being a cost-effective investment option compared to mutual funds, might provide an opportunity to achieve higher returns. Previous studies mentioned above suggest that higher ESG ratings are associated with higher profitability as well as lower risk incidents, which would indicate a potential for abnormal returns. The research by Kumar (2019) and Giese et al. (2019) highlights the argument that higher ESG ratings are associated with lower-risk incidents, higher profitability, and potential increases in returns. These findings suggest that ESG considerations can have a positive impact on financial performance, particularly when reflected in ESG-rated indexes.

However, multiple papers have found inconclusive results concerning the topic. Kanuri (2020) demonstrates that ESG ETFs underperformed the market in terms of both absolute and risk-adjusted measures, indicating that the performance of ESG ETFs can vary over time. Similarly, other studies by Limkriangkrai et al. (2017), Statman and Glushkov (2009), and Kempf and Osthoff (2007) found a positive but not statistically significant effect of ESG ratings on risk-adjusted returns.

It is important to consider the potential causes of the conflicting results. Derwall et al. (2011) and Filbeck et al. (2014) highlight the dynamic nature of socially responsible investing and the complexity of its effects on asset prices and performance. They suggest that the performance of ESG portfolios may be related to specific time frames, investment strategies, or selection biases. Also, the influence of firm characteristics is highlighted as reasons in the literature e.g., financial strength and resource availability on ESG performance. This means that as there are more resources available, newer research will contribute to the old. Hale (2020) notes that the growth of ESG investing accelerated after the 2008 financial crisis, and many ESG ETFs have been launched in the past few years. Furthermore, the timeframe of the studies and the evolution of ESG investing play a role in the observed performance.

The results of this study are consistent with previous studies above. We can draw the conclusion that throughout the period, which ran from 12/2012 to 12/2022, the findings consistently demonstrated that all portfolios, including the S&P 500 ETF Portfolio, Portfolio A (high), and Portfolio B (low), displayed positive alphas, indicating the possibility of abnormal performance. However, the only alpha that was statistically significant across all factor models was Portfolio B's (low) which indicates continuous outperformance outside of the factors considered. The findings do not support hypothesis H1, which states that ESG ETFs with high ESG ratings will see statistically significant anomalous returns when compared to ESG ETFs with low ratings. High ESG ratings and abnormal returns were not significantly correlated, according to the data.

On the contrary portfolio A with high ESG ratings did not regularly show statistically significant abnormal returns. This shows that in the case of passive ESG ETFs, a high ESG rating might not be enough to lead to significant abnormal returns. This indicates that the results are consistent with hypothesis H0, according to which a high ESG rating does not result in statistically significant abnormal returns for ESG ETFs in comparison to ESG ETFs with low ratings. After the data analysis there was no statistically significant abnormal returns found linked to high ESG ratings.

Analyzing the regression model results further reveals that there is still no significant difference in alpha between high and low sustainability ratings. The level of ESG appears to have a minor and nearly neutral impact on abnormal returns. Although Portfolio A (high) and Portfolio B (low) both generated yields that were much lower than the benchmark ETF portfolio, it can be shown that for each regression model, Portfolio A (high) produced a higher yield than Portfolio B (low). Hence, the empirical analysis of this thesis shows that each ETF portfolio displayed positive alphas, indicating the possibility of abnormal performance, yet the ESG portfolio returns are not significant and yielded less than the market benchmark. To support this, previous studies have found that there is a similar connection between the level of ESG scores and the returns yet the returns between high ESG scores and low were not significant enough (Halbritter and Dorfleitner, 2015). That discrepancy could be because of sampling error or the specific selection of ESG criteria. Also, market conditions might affect the results of this study along with limited ESG data. Since there has been significant growth of ESG ETF at the same time with extremely volatile market conditions, it could indicate that we would need to approach the topic from a different point of view e.g., comparing different stages of the market.

Although we believed that ETFs with high ESG ratings will show higher long-term profitability than low ESG rating ETFs. The changes firms make in ESG practices lead to long-term success, improved strategic thinking, value creation, cost advantages through ETFs, reduced financial burden, and increased profitability from attracting a broader group of ESG investors. According to hypothesis H2, long-term profitability is higher for ETFs with strong ESG ratings. The results of this investigation did not offer strong proof in favor of this claim. After the data analysis, there was not found a strong difference in profitability between ETFs with high and low ESG ratings. Hence, the hypothesis H3 will be also rejected. The Sharpe ratio, which offers no substantial difference for Portfolio A and B, further supports this conclusion.

Derwall et al. (2011) highlights the dynamic nature of socially responsible investment (SRI) and its impact on asset prices and performance. They find initial evidence of outperformance for high ESG-rated stocks but suggest diminishing effects over time. Based on the evidence, the time period in this sample for the long-term results has been considerably volatile since the Russian invasion of Ukraine and the global pandemic. As mentioned above that the performance is strongly dependent on the moment of time and development of the ESG ETFs, the long-term results might need a more extended period in the future. The inconsistent returns for environmentally conscious portfolios are indicating a need for further investigation.

Overall, although ESG ratings might influence portfolio performance, the results indicate that ESG rating is not the main driving factor of abnormal returns. Based on all the factor models examined in this study, the other factors have demonstrated varying degrees of influence on portfolio performance including market risk premium (Rm-Rf), SMB, HML, RMW, and CMA. Based on the findings of this paper a direct conclusion could not be made about whether ESG ratings lead to abnormal performance on the performance of ETFs. This is supporting the findings in the previous literature that show contradictory findings on the topic. As Halbritter and Dorfleitner (2015) support that there is a connection between the level of ESG scores and the returns. However, the returns between high ESG scores and low are not significant enough.

The limitation of this study is that the subject is only examined on the U.S. market. Although it is the most prominent market regarding these market trends, there may be problems with this to generalize the results to all markets. The sustainability data for the ETFs was limited and manually picked which could have influenced contradictory results. In addition, although the length of the sample period could have been assumed to be extensive enough, the long-term results might have needed a more extended period. This is due to the significant growth of ESG ETF at the same time with extremely volatile market conditions.

In conclusion, there was found a strong connection between the level of ESG scores and the returns. However, the returns were not significant enough to show the impact of higher ESG performance on financial performance. There was found evidence that ESG ratings might influence portfolio performance. Consequently, the results indicate that ESG rating is not the main driving factor of abnormal returns. Also, when looking at the result, we must consider that the subject is considerably young, and the results might vary strongly depending on the time period. Therefore, empirical research should be continued to obtain more accurate results. Yet, ESG ETFs have become a competitive investment option with the possibility to incorporate personal values into investment decisions. When enough long-term material is available and the market has been able to develop for a longer period of time, it might be possible to find a more clear conclusion for this topic. The research problem is left for further study.

References

Abner, D. J. (2016). *The ETF handbook: How to value and trade exchange-traded funds.* 2nd Edition. New York, United States: John Wiley & Sons Inc.

Arabesque (2020). Why ESG? https://www.arabesque.com/2020/02/04/why-esg/

Auer, B. R. & Schuhmacher, F. (2016). Do socially (ir)responsible investments pay? New evidence from international ESG data. *The Quarterly review of economics and finance*, 59, pp. 51-62. https://doi:10.1016/j.qref.2015.07.002

Bauer, R., Koedijk, K., & Otten, R. (2005). International evidence on ethical mutual fund performance and investment style. *Journal of banking & finance*, 29(7), 1751-1767. https://doi.org/10.1016/j.jbankfin.2004.06.035

Bebchuk, L. A., Cohen, A. & Wang, C. C. (2013). Learning and the disappearing association between governance and returns. *Journal of financial economics*, 108(2), 323-348. https://doi.org/10.1016/j.jfineco.2012.10.004

Ben-David, I., Franzoni, F. & Moussawi, R. (2017). Exchange-Traded Funds. *Annual Review of Financial Economics*, 9(1), 169-189. https://doi.org/10.1146/annurev-financial-110716-032538

Bioy, H. & Lamont, K. (2018). Passive Sustainable Funds: The Global Landscape. *The Journal of Index Investing,* 9(3), 4-17. https://doi.org/10.3905/jii.2018.1.063

BlackRock (2020). Reshaping sustainable investing. https://www.ishares.com/us/literature/whitepaper/reshaping-sustainable-investing-en-us.pdf

Bloomberg (2020). Transforming ESG: A roadmap for creating global standards. https://www.bloomberg.com/professional/blog/transforming-esg-a-roadmap-for-creating-global-standards/

Bloomberg (2019). Conflicting ESG Ratings Are Confusing Sustainable Investors. https://www.bloomberg.com/news/articles/2019-12-11/conflicting-esg-ratings-are-confusing-sustainable-investors

Bodie, Z., Kane, A. & Marcus, A. J. (2018). *Investments.* Eleventh edition. New York: McGraw-Hill Education.

Borgers, A., Derwall, J., Koedijk, K. & Ter Horst, J. (2013). Stakeholder relations and stock returns: On errors in investors' expectations and learning. *Journal of empirical finance,* 22(C), 159-175. https://doi.org/10.1016/j.jempfin.2013.04.003

Buallay, A. (2019). Is sustainability reporting (ESG) associated with performance? Evidence from the European banking sector. *Management of Environmental Quality: An International Journal*, 30(1), pp. 98-115. https://doi:10.1108/MEQ-12-2017-0149

Carhart, M. (1997). On persistence in mutual fund performance. *The Journal of finance (New York)*, 52(1), 57-82. https://doi.org/10.2307/2329556

Chan, K. C. & Chen, N. (1991). Structural and Return Characteristics of Small and Large Firms. *Journal of Finance,* 46(4), 1467-1484. https://doi.org/10.1111/j.1540-6261.1991.tb04626.x

Christensen, D., Serafeim, G., & Sikochi, A. (2019). Why is Corporate Virtue in the Eye of The Beholder? The Case of ESG Ratings. *Working Paper*.

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

Clarkson, P. M., Li, Y., Richardson, G. D. & Vasvari, F. P. (2008). Revisiting the relation between environmental performance and environmental disclosure: An empirical analysis. *Accounting, organizations and society*, 33(4), pp. 303-327. https://doi:10.1016/j.aos.2007.05.003

Cremers, M., Petajisto, A., & Zitzewitz, E. (2012). Should benchmark indices have alpha? Revisiting performance evaluation (No. w18050). *National Bureau of Economic Research.* https://doi.org/10.3386/w18050

Cremers, K. J. M., & Petajisto, A. (2009). How Active Is Your Fund Manager? A New Measure That Predicts Performance*. The Review of financial studies,* 22(9), 3329-3365. https://doi.org/10.1093/rfs/hhp057

Crilly, D. (2012). Faking it or muddling through? understanding decoupling in response to stakeholder pressures*. Academy of Management journal, 55(6), 1429-1448. https://doi.org/10.5465/amj.2010.0697*

Derwall, J., Koedijk, K. & Ter Horst, J. (2011). A tale of values-driven and profit-seeking social investors. *Journal of banking & finance*, 35(8), 2137-2147. https://doi.org/10.1016/j.jbankfin.2011.01.009

Derwall, J., Guenster, N., Bauer, R., & Koedijk, K. (2005). The Eco-Efficiency Premium Puzzle. *Financial analysts journal*, 61(2), 51-63. https://doi.org/10.2469/faj.v61.n2.2716

Dimson, E., Marsh, P. & Staunton, M. (2020). Divergent ESG Ratings. *Journal of Portfolio Management*, 47(1), pp. 75-87. https://doi:10.3905/jpm.2020.1.175

Dorfleitner, G., Halbritter, G. & Nguyen, M. (2015). Measuring the level and risk of corporate responsibility – An empirical comparison of different ESG rating approaches. *Journal of Asset Management,* 16(7), pp. 450-466. https://doi:10.1057/jam.2015.31

Dumitrescu, A., Järvinen, J., & Zakriya, M. (2023). Hidden Gem or Fool's Gold: Can passive ESG ETFs outperform the benchmarks? *International review of financial analysis,* 86, 102540. https://doi.org/10.1016/j.irfa.2023.102540

ECB (2018). Financial Stability Review: Counterparty and liquidity risks in exchange-traded funds. https://www.ecb.europa.eu//pub/financial-stability/fsr/special/html/ecb.fsrart201811_3.en.html#toc1

ETF database (2017). The Hidden Risks and Costs of ETFs. https://etfdb.com/etf-education/hidden-risks-costs-etfs/

Eurosif (2018). EUROPEAN SRI STUDY. http://www.eurosif.org/wp-content/uploads/2018/11/European-SRI-2018-Study.pdf

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

Fama, E. F. & French, K. R. (2004). The Capital Asset Pricing Model: Theory and Evidence. *Journal of Economic Perspectives*, 18(3), pp. 25-46. https://doi:10.1257/0895330042162430

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

Fama, E. & French, K. (1995). Size and book-to-market factors in earnings and returns. *The Journal of Finance,* 50(1), 131. https://doi.org/10.1111/j.1540-6261.1995.tb05169.x

Fama, E. F. & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of financial economics,* 33(1), pp. 3-56. https://doi:10.1016/0304-405X(93)90023-5

Fidelity (2021). ETFs vs. mutual funds: Cost comparison. https://www.fidelity.com/learning-center/investment-products/etf/etfs-cost-comparison

Filbeck, G., Holzhauer, H. & Zhao, X. (2014). Using Social Responsibility Ratings to Outperform the Market: Evidence from Long-Only and Active-Extension Investment Strategies. *Journal of Investing,* 23(1), 79-96,5. https://doi.org/10.3905/joi.2014.23.1.079

Galema, R., Plantinga, A. & Scholtens, B. (2008). The stocks at stake: Return and risk in socially responsible investment. *Journal of banking & finance,* 32(12), pp. 2646-2654. https://doi:10.1016/j.jbankfin.2008.06.002

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

Goss, A. (2009). Corporate social responsibility and financial distress. *Proceedings of the Administrative Sciences Association of Canada*, 29, 1-18.

Halbritter, G. & Dorfleitner, G. (2015). The wages of social responsibility — *where are they? A critical review of ESG investing. Review* of Financial Economics, 26(1), 25-35. https://doi.org/10.1016/j.rfe.2015.03.004

Hale, Jon (2020). Morningstar: Sustainable Equity Funds Are Outperforming in Bear Market. https://www.morningstar.com/articles/972475/sustainable-equity-funds-are-outperforming-in-bear-market

Hill, R., Ainscough, T., Shank, T. & Manullang, D. (2007). Corporate Social Responsibility and Socially Responsible Investing: A Global Perspective. *Journal of Business Ethics,* 70(2), pp. 165-174. https://doi:10.1007/s10551-006-9103-8

Hill, J. M., Nadig, D., & Hougan, M. (2015). A comprehensive guide to exchange-traded funds (ETFs). *CFA Institute Research Foundation*.

HSBC (2019). Why does ESG matter to equity investors? https://www.hsbc.com.cn/en-cn/wealth/insights/esg/2019-03-12/

Hsu, J., Liu, X., Shen, K., Viswanathan, V. & Zhao, Y. (2018). Outperformance through Investing in ESG in Need. *The Journal of Index Investing*, 9(2), 18-26. https://doi.org/10.3905/jii.2018.9.2.018

Huber, B. M. & Comstock, M. (2017). ESG reports and ratings: What they are, why they matter? (SUSTAINABILITY). *Corporate Governance Advisor*, 25(5), 1. https://corpgov.law.harvard.edu/2017/07/27/esg-reports-and-ratings-what-they-are-why-they-matter/

JEGADEESH, N., & TITMAN, S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *The Journal of finance (New York),* 48(1), 65-91. https://doi.org/10.1111/j.1540-6261.1993.tb04702.x

Kanuri, S. (2020). Risk and Return Characteristics of Environmental, Social, and Governance (ESG) Equity ETFs. *The Journal of Index Investing*, 11(2), 66-75. https://doi.org/10.3905/jii.2020.1.092

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

Kumar, R. (2019). ESG: Alpha or Duty? *The Journal of Index Investing*, 9(4), pp. 58-66. https://doi:10.3905/jii.2019.1.066

Limkriangkrai, M., Koh, S. & Durand, R. B. (2017). Environmental, Social, and Governance (ESG) Profiles, Stock Returns, and Financial Policy: Australian Evidence. *International Review of Finance,* 17(3), pp. 461-471. https://doi:10.1111/irfi.12101

Lee, D. D. & Faff, R. W. (2009). Corporate Sustainability Performance and Idiosyncratic Risk: A Global Perspective. *Financial Review*, 44(2), pp. 213-237. https://doi:10.1111/j.1540-6288.2009.00216.x

Lettau, M., & Madhavan, A. (2018). Exchange-traded funds 101 for economists. *Journal of Economic Perspectives,* 32(1), 135-154. https://doi.org/10.1257/jep.32.1.135

Lopatta, K. & Kaspereit, T. (2014). The World Capital Markets' Perception of Sustainability and the Impact of the Financial Crisis. *Journal of Business Ethics,* 122(3), 475-500. https://doi.org/10.1007/s10551-013-1760-9

Lydon, T. (2009). What Are ETFs? ETF trends. https://www.etftrends.com/2009/11/what-are-etfs/

Manescu, C. (2011). Stock Returns in Relation to Environmental, Social and Governance Performance: Mispricing or Compensation for Risk? *Sustainable Development*, 19(2), 95-118. https://doi.org/10.1002/sd.510

Meziani, A. S. (2016). Exchange-traded funds: Investment practices and tactical approaches. https://doi.org/10.1057/978-1-137-39095-0

Mill, G. (2006). The Financial Performance of a Socially Responsible Investment Over Time and a Possible Link with Corporate Social Responsibility. *Journal of Business Ethics*, 63(2), 131-148. https://doi.org/10.1007/s10551-005-2410-7

MSCI (2023). ESG Ratings. https://www.msci.com/our-solutions/esg-investing/esg-ratings

MSCI (2020a). Deconstructing ESG Ratings Performance - Risk and Return for E, S and G by Time Horizon, Sector and Weighting. https://www.msci.com/documents/10199/7b77de78-0c6e-0a45-f4dd-e65025552bae

MSCI (2020b). ESG 101: What is ESG? https://www.msci.com/what-is-esg

Nasdaq (2021). An ETF to Watch As Pandemic Spurs Demand for Global ESG. https://www.nasdaq.com/articles/an-etf-to-watch-as-pandemic-spurs-demand-for-global-esg-2021-09-09

Nofsinger, J., & Varma, A. (2014). Socially responsible funds and market crises. *Journal of banking & finance,* 48, 180-193. https://doi.org/10.1016/j.jbankfin.2013.12.016

Puttonen, V. and Repo, E., 2007. *Miten sijoitan rahastoihin*. 4th edition. Helsinki: WSOYpro.

Ramaswamy, S. 2011. Market structures and systemic risks of exchange-traded funds. *Working Paper*, Bank Int. Settl., Basel, Switz.

Sarkar, D., Datta, R., Mukherjee, A. & Hannigan, R. (2016). An integrated approach to environmental management. https://doi.org/10.1002/9781118744406

Sharfman, M. P. & Fernando, C. S. (2008). Environmental risk management and the cost of capital. *Strategic Management Journal,* 29(6), pp. 569-592. https://doi:10.1002/smj.678

Sharpe, W. (1994). The Sharpe Ratio. *Journal of Portfolio Management*, 21(1), p. 49. https://doi:10.3905/jpm.1994.409501

Sharpe, W. (1975). ADJUSTING FOR RISK IN PORTFOLIO PERFORMANCE MEASUREMENT. *Journal of Portfolio Management,* 1(2), p. 29. https://doi:10.3905/jpm.1975.408513

Sharpe, W. F. (1964). CAPITAL ASSET PRICES: A THEORY OF MARKET EQUILIBRIUM UNDER CONDITIONS OF RISK*. *Journal of Finance,* 19(3), pp. 425-442. https://doi:10.1111/j.1540-6261.1964.tb02865.x

SSGA (2020a). SPDR® ETFs Throughout the COVID-19 Crisis. https://www.ssga.com/library-content/pdfs/etf/spdr-amid-the-covid19-crisis.pdf

SSGA (2020b). SPDR® ESG investing tipping point to turning point https://www.ssga.com/library-content/pdfs/etf/spdr-esg-investing-tipping-point-to-turning-point.pdf

Statista (2022). Development of assets of global exchange traded funds (ETFs) from 2003 to 2021. https://www.statista.com/statistics/224579/worldwide-etf-assets-under-management-since-1997/

Statman, M. & Glushkov, D. (2009). The wages of social responsibility. *Financial Analysts Journal,* 65(4), p. 33. https://doi:10.2469/faj.v65.n4.5

Stevenson,D. and Tuckwell, D. (2019). *The Ultimate ETF Guidebook.* Harriman House Publising.

The Asset (2022). ESG ETFs attract US$7 billion net inflow in March. https://www.theasset.com/article-esg/46603/esg-etfs-attract-us-7-billion-net-inflow-in-march

Trackinsight (2022). Introducing the three largest ESG ETFs. https://www.trackinsight.com/en/etf-news/introducing-three-largest-esg-etfs

UN PRI (2020). What are the Principles for Responsible Investment? https://www.unpri.org/pri/what-are-the-principles-for-responsible-investment

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