UNIVERSITY OF LODZ

FACULTY OF ECONOMICS AND SOCIOLOGY

DOCTORAL SCHOOL OF SOCIAL SCIENCES

DISCIPLINE: ECONOMICS AND FINANCE

Katarzyna Daniluk

ESG Investing and Passive Fund Management

Tracking Ability of Passive ESG Equity Exchange-Traded Funds

Doctoral dissertation
prepared under the academic supervision of
Tomasz Miziołek, Associate Professor
at the University of Lodz,
and Dr Ewa Feder-Sempach,
at the Department of International Finance
and Investment.

Lodz 2025

TABLE OF CONTENTS

ABSTRACT 4

INTRODUCTION 5

CHAPTER 1 PASSIVE PORTFOLIO MANAGEMENT 12

1. 1 Introduction 12
1. 2 Theoretical Background 13
1. 2. 1 Fundamentals of Portfolio Theory 13
1. 2. 2 Capital Market Models 18
1. 2. 3 Efficient Market Hypothesis 25
1. 3 Active vs. Passive Asset Management 30
1. 3. 1 Key Research on Active vs. Passive Investing 31
1. 3. 2 Significance of Costs 37
1. 3. 3 The Rise and Controversies of Passive Investing 40
1. 4 Indices as Benchmarks in Passive Portfolio Management 42
1. 4. 1 Index Functions 42
1. 4. 2 Benchmark Index Construction 44
1. 4. 3 Index Industry 48
1. 5 Summary 51

CHAPTER 2 ESG INVESTING 52

2. 1 Introduction 52
2. 2 ESG Terminology 53
2.2.1 Defining ESG investing 53
2.2.2 ESG Strategies 59
2. 3 ESG Dimensions 60
2. 3. 1 Environmental Factors 61
2. 3. 2 Social Factors 62
2. 3. 3 Governance Factors 66
2. 4 ESG Initiatives 68
2. 4. 1 United Nations Initiatives 69
2. 4. 2 Reporting Initiatives 72
2. 5 Investment Performance and Risk of ESG Investments 74
2. 5. 1 Theoretical Considerations 74
2. 5. 2 Sin Stocks Performance 76
2. 5. 3 ESG and Risk-adjusted returns 77

2.6 Passive ESG Approach 83
2.7 Summary 88
CHAPTER 3 Benchmark Replication in ETFs 89
3. 1 Introduction 89
3. 2 Essentials of Exchange-Traded Funds 90
3. 2. 1 Operation of Exchange-Traded Funds 91
3. 2. 2 Methods of Index Replication 94
3. 2. 3 Valuation Mechanisms and Price Dynamics 97
3. 2. 4 ETF Liquidity 99
3. 3 Advantages of Exchange-Traded Funds 101
3. 3. 1 Portfolio Diversification 101
3. 3. 2 Cost Efficiency and Trading Flexibility 103
3. 3. 3 Transparency 104
3. 3. 4 Versatility 107
3. 4 Tracking Performance of ETFs 109
3. 4. 1 Metrics for Evaluating Tracking Accuracy 110
3. 4. 2 Research on ETFs Tracking Ability 112
3. 4. 3 Factors Influencing Tracking Performance of ETFs 114
3. 5 Summary 118

CHAPTER 4 TRACKING ABILITY OF PASSIVE ESG EQUITY ETFs – EVIDENCE FROM EUROPEAN MARKET 119

4. 1 Introduction 119
4. 2 Research Design and Methods 119
4. 3 Research Sample Characteristics 129
4. 4 Empirical Results and Discussion 137
4. 4. 1 Comparison of Tracking Errors in ESG and Non-ESG ETFs 137
4. 4. 2 Modelling Tracking Error of ETFs 140
4. 5 Discussion 158
4. 6 Summary 162

CONCLUSION 163

BIBLIOGRAPHY 167

LIST OF TABLES 181

LIST OF FIGURES 183

ABSTRACT

The rapid growth of passive investing and the increasing emphasis on Environmental, Social, and Governance (ESG) criteria have intersected to form a new class of investment products: passive ESG equity Exchange-Traded Funds (ETFs). While these funds aim to offer both market exposure and sustainable investment features, concerns remain regarding their ability to accurately replicate benchmark indices.

The primary objective of this dissertation is to evaluate and compare the tracking ability of passive ESG equity ETFs listed on European exchanges, in relation to their non-ESG counterparts, and to identify the key factors affecting replication quality. The study examines 48 ESG and 86 non-ESG ETFs throughout 2021 to 2024 using comparative analysis and dynamic panel (GMM) estimation. The replication performance is measured by tracking error through three standard metrics. The study explores how fund characteristics such as total expense ratio (TER), assets under management (AUM), age, benchmark volatility, and replication method affect the tracking performance of ETFs.

The research shows that ESG integration does not negatively affect and may even improve the replication performance of passive equity ETFs. The constructed dynamic panel models show that tracking errors of both ESG and non-ESG ETFs persist strongly from one period to the next. The TER and benchmark volatility increase tracking errors, but AUM reduces tracking errors only in ESG ETFs. The tracking performance of both groups remains unaffected by fund age.

The Chow test established that fund size and cost affect tracking error differently between the two groups. However, the extended panel dynamic GMM model that included interaction terms did not confirm stable differences between groups. The model indicated that fundamental factors that determine tracking error performance remain largely the same for passive ESG equity ETFs and non-ESG ETFs listed on European exchanges.

The dissertation adds to academic knowledge about ETF replication efficiency while offering practical insights to improve passive ESG investment product management. The study confirms that ESG ETFs function as investment vehicles that unite sustainability with good replication performance.

INTRODUCTION

Passive portfolio management and Environmental, Social, and Governance (ESG) investing have become two trends shaping global capital markets since the early 2000s and 2010s, respectively. The Efficient Market Hypothesis (Fama, 1970) supports passive investing based on its premise that financial markets reflect all accessible information. Consequently, attempting to outperform the market is considered less effective over the long term than adopting a strategy that seeks to replicate overall market returns. Index tracking serves as a fundamental mechanism that enables passive investing through the creation of portfolios that replicate market index performance (Betides et al., 2018). However, market indices exist as theoretical frameworks, like "paper portfolios" (Roll, 1992) that investors cannot directly invest in. Investors use tangible investment products such as Exchange-Traded Funds (ETFs) to access market indices. These funds strive to replicate index performance while dealing with real-world constraints, including trading costs, liquidity frictions, and rebalancing requirements (Johnson, 2009; Frino & Gallagher, 2001).

Over the past two decades, ETFs have become the dominant instrument for implementing passive strategies, offering investors low-cost, transparent, and flexible exposure to market benchmarks (Agapova, 2011; Rompotis, 2011). The replication efficiency of ETFs—measured by tracking error, i.e. the deviation between ETF returns and benchmark returns— is a critical tracking performance metric. A well-managed passive ETF should minimize tracking error, thus preserving the core value proposition of passive investing. However, the increasing incorporation of Environmental, Social, and Governance (ESG) criteria into index construction has introduced new structural complexities into this framework, which may change traditional patterns of replication efficiency.

Although ESG factors were originally linked to active investment approaches, they are now increasingly incorporated into the design of benchmark indices (Boffo & Patalano, 2020). However, the implementation of ESG filters within index methodologies creates fundamental differences between ESG benchmarks and traditional benchmarks. The exclusion of companies involved in controversial sectors and those with weak ESG performance leads to systematic changes in index composition and weighting schemes (MSCI Inc., 2023). The modifications improve the sustainability profiles of investment portfolios but may also affect the replication accuracy of passive funds that track ESG indices.

The research problem of this dissertation concerns the impact of ESG integration on the tracking ability of passive equity ETFs. It focuses on differences in tracking error between

ESG and non-ESG ETFs listed on European exchanges and examines how selected factors influence the replication accuracy of these funds. The study is limited to traditional passive investing through broad-market ETFs encompassing ESG and non-ESG funds. It excludes ETFs that, despite passive structures, incorporate active strategies, such as smart-beta, factor-based, leveraged, inverse, or currency-hedged products. The tracking ability shows the degree to which an ETF mirrors its benchmark. Effective index replication is crucial both from the perspective of the investor and the manager of passive funds, as the good tracking ability of an ETF determines the efficiency and quality of a passive investment product.

The literature on the tracking ability of ETFs is well-developed and covers the US market (Elton et al., 2002; Poterba and Shoven, 2002), other developed markets (Gallagher and Segara, 2006; Rompotis, 2008), as well as emerging markets (Chu, 2011; Khan et al., 2015; Miziołek and Feder-Sempach, 2019). Primarily, the tracking performance of ETFs depends on total expense ratio (TER) (Frino and Gallagher, 2001; Rompotis, 2009; Blitz et al., 2012), fund size (Chu, 2011), and market volatility (Qadan and Yagil, 2012; Drenovak et al., 2014).

Even though the market of ESG ETFs is growing rapidly, with 1,546 funds managing over USD 645.21 billion in November 2024 (ETFGI, 2024), the research on their replication performance remains scarce. The majority of existing research focuses on ESG's effects on returns (Winegarden, 2019; Plagge; Grim, 2020; Milonas et al., 2022), risk mitigation (Kanuri, 2022; Huang, 2024) and portfolio diversification (Cornell & Damodaran, 2020; Pedersen et al., 2021; Pastor et al., 2021) instead of tracking precision (Nguyen, 2023). The impact of ESG factors on ETF tracking errors has not been thoroughly studied yet. This lack of research creates uncertainty for investors who are looking for sustainable and low-cost passive investment options.

The tracking performance of passive ESG ETFs remains a concern, making them less popular among investors. Many choose non-ESG ETFs because they have a longer history of results and lower costs (Dumitrescu et al., 2023). Creating ESG indices is more complicated, as it requires both financial and ESG-related data, which increases the complexity and may lead to higher licensing fees for passive funds. On the other hand, the adoption of artificial intelligence (AI) for data processing has been reported by 42% of ESG data providers according to Substantive Research (2024). Then, the implementation of the Sustainable Finance Disclosure Regulation (SFDR) in the European Union, together with the so-called "price war" among ETFs, has lowered ESG investing costs. The European Securities and Markets Authority (ESMA, 2025) reports that fees for ESG funds in Europe match fees for non-ESG funds.

Soupe and Kovarcik (2024) discovered that equity funds with ESG objectives tend to achieve higher returns, better tracking, and lower tracking errors. Nguyen (2023) analyzed 126 U.S. ESG ETFs from January 2019 to March 2022 and found that these funds maintained solid tracking performance even during the COVID-19 market turmoil. However, the study was limited to specific countries and timeframes and did not investigate the broader factors influencing tracking performance. In contrast, this study aims to address these limitations by exploring the tracking efficiency of ESG ETFs listed on European exchanges throughout an extended period relative to non-ESG ETFs, while using fund-level characteristics and market conditions in the evaluation framework.

The primary objective of this dissertation is to evaluate and compare the tracking ability of passive ESG equity ETFs listed on European exchanges, in relation to their non-ESG counterparts, and to identify the key factors affecting replication quality. The specific objectives are as follows:

- To compare the quality of index replication between ESG and non-ESG passive equity ETFs listed on European exchanges.
- To examine the impact of fund total expense ratio (TER), assets under management (AUM), age, and benchmark volatility on the index replication quality of passive broad market ESG equity ETFs listed on European exchanges, and their non-ESG counterparts.
- To assess whether the impact of these factors on tracking error differs between passive broad market ESG equity ETFs listed on European exchanges and their non-ESG counterparts.

According to Avramov et al. (2022), ESG integration leads to more stable cash flows and better risk management. On the other hand, Pedersen et al. (2021) state that excluding companies with low ESG ratings can limit portfolio diversification. The study conducted by Ling et al. (2023) revealed that the structural differences between ESG and traditional indices may affect the consistency of replication. Accordingly, the first hypothesis of this study states:

H1: Passive ESG equity ETFs listed on European exchanges exhibit significantly different tracking errors compared to their non-ESG counterparts.

Next, ESG ETFs have unique operational characteristics. The asset pools managed by these funds remain relatively small (Fund Selector Asia, 2024), and ESG screening leads to higher licensing expenses (BFinance, 2024). Moreover, quality biases may emerge as the result of specific factor tilts (Ascioglu and Saatcioglu, 2024). The distinct features of these funds may change how traditional factors influence their replication accuracy. Thus, the second general hypothesis is:

H2: The determinants of tracking error differ between passive ESG equity ETFs listed on European exchanges and their non-ESG counterparts.

To better understand these patterns, this dissertation tests five specific hypotheses. In line with prior research, which indicated the autoregressive pattern of ETF tracking errors (DeFusco et al., 2011; Ivanov, 2015), the first specific hypothesis states:

HS1: Tracking errors of both ESG and non-ESG passive equity ETFs listed on European exchanges exhibit an autoregressive pattern.

Second, the literature consistently confirms that higher fund costs reduce replication quality. Authors such as Frino and Gallagher (2001) and Blitz et al. (2012) found a positive relationship between TER and tracking error levels. Accordingly:

HS2: There is a positive relationship between the total expense ratio (TER) and the tracking error of both ESG and non-ESG passive equity ETFs listed on European exchanges.

Third, previous research evidenced that larger funds tend to achieve lower tracking errors since they benefit from economies of scale (Chu, 2011; Dorocáková, 2017). Thus, the next hypothesis is:

HS3: There is a negative relationship between assets under management (AUM) and the tracking error of both ESG and non-ESG passive equity ETFs listed on European exchanges.

Next, according to Rompotis (2011), older ETFs display lower tracking errors because their operational maturity contributes to cost efficiencies and improved liquidity. Consequently:

HS4: There is a negative relationship between fund age and the tracking error of both ESG and non-ESG passive equity ETFs listed on European exchanges.

Finally, studies by Qadan and Yagil (2012) and Drenovak et al. (2014) suggested that benchmark volatility increases tracking errors as a consequence of increased rebalancing needs and higher transaction costs. Therefore, the final specific hypothesis states:

HS5: There is a positive relationship between benchmark volatility and the tracking error of both ESG and non-ESG passive equity ETFs listed on European exchanges.

To meet the objectives of the dissertation and test the hypotheses, the study uses a critical review of the literature, descriptive statistics, comparative analysis, and panel linear regression models. The literature review looks at key theories related to passive investing and ESG, explains how ETFs work, and explores what tracking ability is, how it is measured, and what factors influence it in passively managed ETFs. Then, descriptive statistical methods pertain to evaluating the tracking error indicators of 48 ESG ETFs and 86 non-ESG ETFs using three methods: tracking error calculated as the absolute return difference between the fund and the benchmark index (TE_1), tracking error calculated as the standard deviation of these return differences (TE_2), and tracking error calculated as the standard deviation of the residuals from a regression of the ETF return on the benchmark return (TE_3). Determinants of tracking error analyzed included assets under management (AUM), total expense ratio (TER), age of the ETF, benchmark volatility, and index replication method. Next, comparative analysis entails comparing the level and determinants of tracking ability measures between ESG and non-ESG ETFs. The significance of differences in the TE_1, TE_2, and TE_3 indicators and tracking error determinants between ESG and non-ESG ETFs was assessed using Wilcoxon rank-sum tests. Finally, panel linear regression models are used to analyze and compare the impact of AUM, TER, ETF age, benchmark volatility, and index replication method on the tracking ability of passive ESG equity ETFs listed on European exchanges and their non-ESG counterparts.

The primary research dilemmas and limitations stemmed from the research sample and study period selection. Eventually, the study focuses on the European ETF market due to its dominance in the global ESG ETF industry and the supportive regulatory environment established by the EU, such as the SFDR and EU Taxonomy Regulation, which promote transparency and accountability in ESG investing. The research covers the period from January 2021 to June 2024 to ensure full and reliable data. Before 2021, there were too few ESG ETFs to conduct a solid analysis. This period reflects rapid growth in the ESG ETF market (ETFGI, 2025) and captures the effects of the introduction of the SFDR regulation in 2021, which improved transparency and limited greenwashing.

The dissertation contributes to the existing financial literature as follows. Firstly, using multiple methods (absolute return difference, standard deviation of return differences, and regression residuals), the study provides a detailed comparative analysis of the tracking ability of ESG and non-ESG ETFs listed in European exchanges. By comparing the tracking errors of ESG and non-ESG ETFs, this study provides useful insights into how well each type follows its index. This comparison makes the research more practical for investors, fund managers, and researchers, as it shows how ESG ETFs perform compared to more established non-ESG ETFs. This helps set expectations and benchmarks for ESG funds. Including both types of ETFs also provides a fuller picture of the passive ETF market, making the results relevant not only for market participants but also for regulators and supervisors.

Then, the research quantifies the impact of various factors like total expense ratio, assets under management, fund age, benchmark volatility, and index replication method on the tracking ability of both ESG and non-ESG ETFs. Comparing these determinants between

ESG and non-ESG ETFs delivers valuable information on whether the same factors influence both types of funds similarly or if unique factors affect ESG ETFs. Fund managers can use the findings to more effectively design and manage ESG ETFs, which may improve their tracking performance and make them more appealing to investors.

This dissertation helps raise awareness about how well ESG ETFs follow their benchmarks, which can correct false beliefs and support wider acceptance of sustainable investing. By presenting clear conclusions about tracking ability and the factors that influence it for both ESG and non-ESG ETFs, the study gives investors useful information to make more confident and informed decisions. Providing empirical evidence that ESG ETFs can accurately replicate their benchmarks helps to build investor trust. This is important for addressing doubts about whether passive ESG investing can be financially effective and for supporting the growth of this investment area.

The dissertation consists of four chapters, an introduction, and a conclusion. The first chapter concentrates on passive portfolio management. It explains the basics of portfolio theory, capital market models, and the efficient market hypothesis. The next section reviews research on active and passive investing, analyses the importance of costs in the success of passive investing, and explores the observed shift towards passive investing. The last part of the chapter addresses indices representing benchmarks for passively managed funds. It provides a broader background to the financial market indices application, explains how index benchmarks are constructed, and characterizes the index industry.

The second chapter deals with ESG investing. The first part of the chapter attempts to systematise the terminology, concepts, and strategies associated with ESG investing. Subsequently, it synthesises ideas lying behind the ESG acronym and characterises ESG dimensions. This is followed by an overview of the key ESG initiatives that have driven the development of the ESG industry. The consecutive subchapter considers the theoretical rationale behind the performance and risk of ESG investments and reviews the literature on the returns of ESG investments. The final part of the chapter focuses on the characteristics of the passive ESG investment approach underpinning passive ESG ETFs and reviews the literature on this topic.

The third chapter of the dissertation focuses on ETFs and their tracking ability. The first part of the chapter aims to systematise and characterise the fundamental aspects of the ETFs' operation. This includes exchange-traded funds functioning, the index replication methods used by passively managed funds, ETF valuation and pricing issues, and their liquidity. The next part of the chapter is a literature review of the benefits of ETFs, i.e. diversification, low costs, trading flexibility, transparency, and versatility of this instrument. The last section reviews the literature on the tracking ability of ETFs. It presents different methods of measuring tracking ability, existing research on the issue, and determinants of tracking ability.

The final chapter is devoted to the characteristics of the empirical analysis results of the tracking ability of passive ESG equity ETFs listed in European exchanges and their determinants. The following subchapters describe the research sample, methods, empirical results, discussion, and conclusions.

In conclusion, this dissertation contributes to financial research by examining the tracking performance of passive ESG ETFs, with a particular focus on the European ETF market. The research explores the impact of the ESG factor on the ETF tracking performance, as well as its determinants. It evaluates multiple factors, including total expense ratio, assets under management, fund age, and benchmark volatility, to determine both distinctive and shared patterns of ESG and non-ESG ETFs tracking performance. Therefore, the research develops academic knowledge on ETF replication efficiency and provides practical insights to enhance the management of passive ESG investment products.

CHAPTER 1 PASSIVE PORTFOLIO MANAGEMENT

1. 1 Introduction

Passive ESG equity ETFs that constitute the principal interest of the thesis represent an extension of the previous innovations in the financial markets. Collective investment dates back to the second half of the 18th century when the first institution resembling today's collective investment funds was established in the Netherlands, for years investing through funds was reserved only for the wealthiest members of the society. In 1774, a Dutch businessman and broker, Abraham van Ketwich, founded an investment trust called “Eendragt Maakt Magt”, which stands for “unity gives strength”. Its aim was to enable minor investors with small financial surpluses to better diversify their investment portfolio. In fact, this basic purpose of funds is still valid today. Ketwich's idea was based on reducing the risk of individual investors by allocating their funds to different asset classes. Then, the development of the mutual fund market in the 18th century was a gradual process, during which merchants and brokers learnt how to expand the range of investment opportunities for the public. The introduction of financial innovations such as securitisation (converting illiquid debt into securities that can be traded on financial markets) and substitution (converting the securities into individual assets or parts of a portfolio to facilitate their trading) has overcome previous barriers and allowed retail investors to earn money on foreign stock exchanges (Rouwenhorst, 2004: 1–2).

The creation of the first open-end fund was of great importance for the development of the mutual fund market, and typically ETFs operate in this form. While previously all funds operated as closed-end funds, the Massachusetts Investors Trust, which was established in Boston in 1924, had the form of an open-end fund. Nowadays, most funds operate as open-end funds, so a fund can create and redeem its shares as required. This is essential from the perspective of the flexibility of buying and selling units and has allowed mutual funds to become widespread.

However, for decades, all mutual funds have actively selected securities for portfolios to generate the highest possible return for participants, which was associated with significant costs. This has started to change only when another key innovation in the financial markets came – passive investing, which has completely changed the approach to portfolio management and whose role has been growing steadily over the years. The first index fund, known as the First Index Investment Trust, was created by John C. Bogle, the founder of Vanguard Group, and was launched in 1976. The concept of passive investing is based on passively tracking the performance and risk of the broad market. Instead of actively selecting assets for the portfolio, it involves mimicking the results of a broad market and accepting a return close to the average return of instruments listed in a specific market.

Essentially, passive portfolio management provides the foundation for understanding the construction and management of passive investment vehicles like ETFs, which are prime subject of the thesis. As passive ETFs are designed to replicate the performance of their underlying index, the concept of passive portfolio management is directly applicable to them. By exploring passive portfolio management, it is possible to gain insight into the strategies, methodologies, and challenges associated with achieving accurate tracking in ETFs. Finally, to understand how passive funds function, it is necessary to explain which financial theories gave rise to the idea of passive investment, what justifies this investment approach, and what it is based on. Thus, to provide some contextual understanding, the following chapter aims to explore the advances in financial market theory which gave rise to the passive investing, as well as some research proving the rationale behind passive investing and the essence of indices which constitute the base for passive portfolio management.

1. 2 Theoretical Background

Understanding the idea of passive portfolio management demands getting familiarised with the key achievements of finance theory concerning investments. Thus, the following part of the thesis is devoted to presenting the crucial aspects of portfolio theory, capital market models, and efficient market hypothesis which significantly contributed to the development of passive investing.

1. 2. 1 Fundamentals of Portfolio Theory

As the core of any investment fund is to invest in a portfolio of assets rather than a single instrument, discussion of passively managed funds should begin with the basic achievements of portfolio analysis. However, before proceeding to discuss the fundamentals of portfolio theory, it may be appropriate to mention two primary terms describing any investment, which are return and risk. While the definition of return, which is simply a change in the price of a financial instrument, requires no explanation, the essence of risk in financial investments is not so obvious and thus, might be misunderstood. The most famous definition of risk was provided by Knight (1921), who made a distinction between risk and uncertainty. He stated that risk is the possibility of alternative outcomes whose probabilities are measurable, and uncertainty, in contrast, could not be measured. Indeed, if the risk depicted all the randomness in financial markets, professional financial institutions would be able to price and sell investment products, which rely solely on quantifiable phenomena. Therefore, one should be aware that, in addition, investing encompasses uncertainty, which creates constraints that these institutions may not be able to calculate and predict. Then, Ellsberg (1961) proposed an even more precise definition of uncertainty based on the concept of probability, indicating that an event is uncertain if it has an unknown probability. This distinction is crucial from a portfolio theory perspective, as it is risk, not uncertainty, that is the focus of portfolio analysis.

Another prominent issue that laid the foundations for the portfolio theory is the classical expected utility theory developed by John von Neumann and Oskar Morgenstern (1944). According to this theory, the behaviour of individuals is the result of their desire to maximize the expected value of the individual utility function. Relating to investing and portfolio theory, this means that investors do not choose the highest possible expected value, understood as the highest rate of return, but the highest expected value of utility. Consequently, the utility curve looks different for risk-takers, risk-neutral, and risk-averse investors.

The principal relation between risk and return is positive and linear and might be shown in a simple graph known as the security market line (SML). As a rule, with an increase in the risk of an asset, the required rate of return rises. The SML presents possible risk-return combinations for all risky assets in the capital market at a given time (see Figure 1). The risk-free rate (RFR) standing for investment with no uncertainty and risk is also marked.

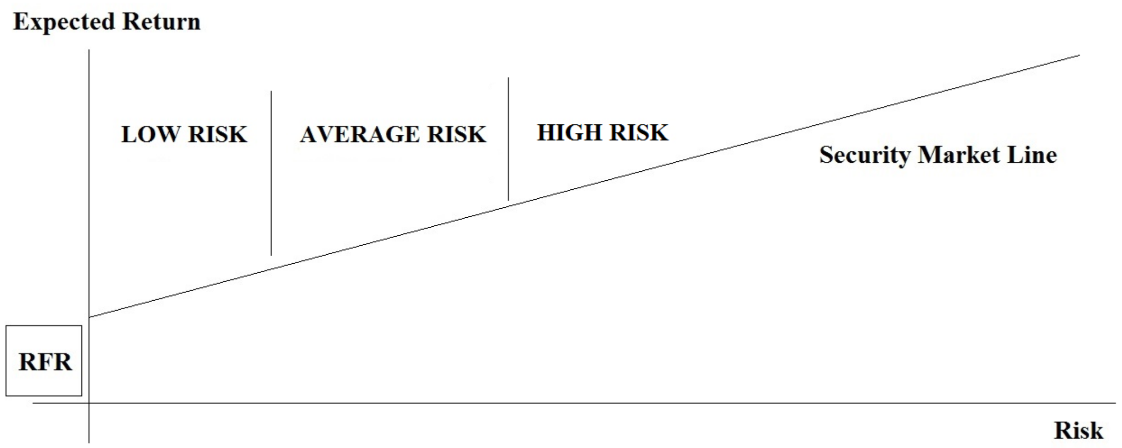

Figure 1. Relationship between risk and return

(Source: Own elaboration based on Elton et al., 2014: 296)

Depending on investors risk preferences, they choose an investment with low, average or high risk exposure. As the risk rises, investors are expecting higher and higher returns.

After explaining the basic dependencies between return and risk, it is time to focus on portfolio theory. The cornerstone of the portfolio theory is the identification of the optimal structure of capital investments. To determine the optimal composition of a portfolio (a set of assets with corresponding shares in the portfolio), two previously mentioned investment characteristics are analysed: risk and return. The foundations of modern portfolio theory derive from the work of Harry Markowitz. It was he who first said that investment decisions are based not only on the expected rate of return but also on considering the investment risk. His article *Portfolio Selection* (1952), followed by the monograph *Portfolio Selection: Efficient Diversification of Investments* (1959) provided the basis for constructing an optimal investment portfolio. For his contributions to the field, in 1990, he was awarded the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel. Markowitz's considerations were centred on the question of how to select and combine assets in a portfolio in such a way that would give the investor the highest possible rate of return at a given level of risk or minimise risk for a given level of return.

Markowitz's model was based on several assumptions concerning investor behaviour:

1. Investors consider each investment to be represented by a probability distribution of expected returns over some holding period.
2. Investors maximize one-period expected utility, and their utility curves demonstrate diminishing marginal utility of wealth.
3. Investors estimate the risk of the portfolio on the basis of the variability of expected returns.
4. Investors make decisions only on the basis of expected return and risk, so their utility curves are a function of expected return and expected variance (or standard deviation) of returns.
5. For a given risk level, investors prefer higher returns to lower returns and, similarly, for a given level of expected return, they prefer less risk to more risk.

According to the Modern Portfolio Theory (MPT), under the above assumptions, a single asset or portfolio of assets is considered to be efficient if there is no other asset or portfolio of assets, which offers higher expected return with the same (or lower) risk or lower risk with the same (or higher) expected return.

The considerations regarding the characteristics of individual investments presented earlier can be then applied to a multi-component investment portfolio. The expected return for a multi-asset portfolio with *n* components is a weighted average of the individual components of the portfolio, which may be written as follows:

$$E(R_P) = \sum_{i=1}^{n} W_i \, R_i;$$

where:

$E(R_P)$ – The expected rate of return of the portfolio,

$W_i -$ The weight of asset *i* in the portfolio p (the percent),

$R_i -$ The expected rate of return for asset *i*.

To calculate the risk of a multi-asset portfolio, it is necessary not only to know the variances (or standard deviations) of the individual assets but also the covariances between pairs of individual securities in the portfolio.

According to the Modern Portfolio Theory (MPT), risk might be calculated as the standard deviation of returns. The general formula for the portfolio standard deviation can be written as follows:

$$\sigma_{portolio} = \sqrt{\sum_{i=1}^{n} w_i^2 \sigma_i^2 + \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j Cov_{i,j}}\,;$$

where:

$\sigma_{portolio} -$ The standard deviation of the portfolio,

$w_i -$ The weights of the individual assets in the portfolio, where weights are determined by the proportion of value in the portfolio,

$\sigma_i^2 -$ The variance of rates of return for asset *i*,

$Cov_{i,j} -$ The covariance between rates of return for assets *i* and j.

Markowitz explained that the standard deviation for a portfolio is not just the weighted average of the standard deviations of individual assets. The portfolio standard deviation encompasses both the weighted average of the individual variances of assets in the portfolio and the weighted covariances between all pairs of individual assets. Therefore, when selecting assets for the portfolio, investors should pay attention not only to the returns and risks of individual instruments but also to the correlations between these returns. From the view of portfolio management practice, it means that by adding to the portfolio more financial instruments with a similar return but low or negative correlation, it is possible to reduce the risk of the portfolio without lowering the return. Moreover, when there is a large number of securities in the portfolio, the portfolio standard deviation formula reduces to the sum of weighted covariances. The traditional low-cost passive investing approach draws on the Modern Portfolio Theory by investing in a bunch of securities that deliver a diversified portfolio.

Another pivotal issue in portfolio theory is identifying the set of investment opportunities that provide the most optimal relationship between return and risk while taking into account the individual preferences of an investor. The curve that illustrates all the possible effective combinations of assets is called the efficient frontier. The efficient frontier depicts a set of portfolios that have the maximum rate of return for every given level of risk or the minimum risk for every level of return (see Figure 2). It represents all the portfolios that any rational, risk-averse investor would choose. As shown below, portfolios A and B are efficient, while portfolio C is not.

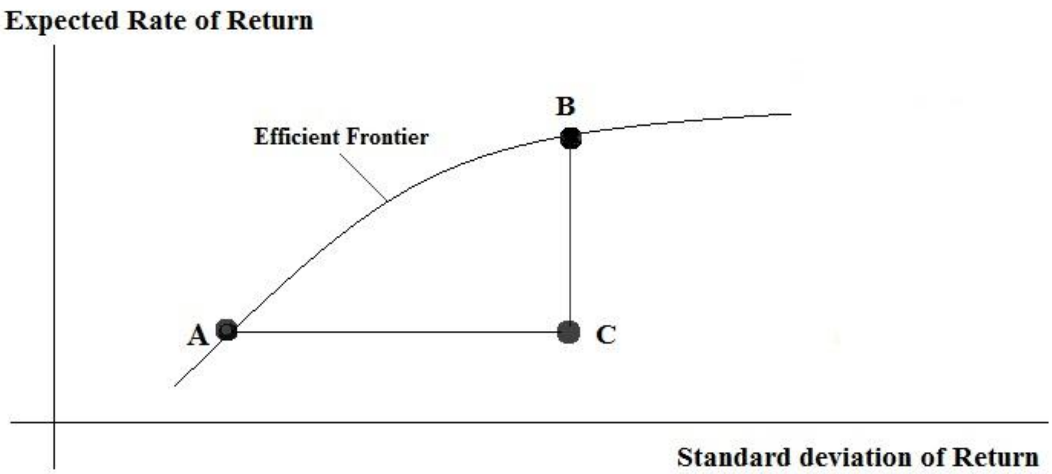

Figure 2. The efficient frontier for alternative portfolios
(Source: Own elaboration based on Reilly and Brown, 2012: 199)

The efficient frontier is often graphically represented, enabling investors to visually evaluate the risk-return profile of various portfolios. Undoubtedly, it is a convenient solution for investors as it can facilitate communication, decision-making, and understanding of trade-offs among different investment opportunities. However, as the theory is years old, it does not include additional investors’ preferences combined with non-financial considerations, which might be perceived as a crucial limitation for ESG investors. As a result of the extension of Markowitz's efficient frontier, Pedersen et al. (2021) proposed an ESG efficient frontier, which will be a subject of debate in the chapter about ESG investing.

What is more, a significant drawback of the efficient frontier is that it is typically constructed based on a single period's data and does not consider the dynamic nature of investment decisions. In the real financial world full of market turbulences and changing market conditions, investor risk preferences change over time. This means that any time an investor alters his preferences, he or she needs to construct a new efficient frontier.

Some extension and improvement of the Modern Portfolio Theory was the separation theorem formulated by James Tobin (1958). What was new in his theory was the consideration of risk-free instruments. According to Tobin, the process of selecting assets for a portfolio can be divided into two stages. The first one is to find an efficient portfolio of risky assets following the Modern Portfolio Theory. Then, the second one is to determine the optimum fraction of the portfolio consisting of risky assets and the fraction consisting of a risk-free asset, depending on the investor's level of risk aversion. A risk-free asset is an asset that has a certain future rate of return, for instance, treasury bonds or notes. The expected rate of return of a portfolio according to the Tobin model can be expressed by the following formula:

$$E\left(R_p\right) = wE(R_m) + (1-w)R_f;$$

where:

$E\left(R_p\right) -$ The expected rate of return of the portfolio composed of risky and risk-free assets,

$w -$ The weight of risky assets in the portfolio,

$E(R_m) -$ The expected rate of return on risky assets,

$R_f -$ The rate of return on risk-free assets.

The foundations of modern portfolio theory developed by Markowitz and Tobin enabled the further development of research on investment portfolios and capital market behaviour. The following part of the paper will present other pivotal contributions to the passive investing approach.

1. 2. 2 Capital Market Models

Despite the Markowitz (1952) and Tobin (1958) models being revolutionary and of great importance, they were challenging to apply to practice. Not only did they require the calculation of returns of every security and their standard deviations, but also correlation coefficients between each pair of individual securities in the portfolio, which turned out to be highly time-consuming. Therefore, efforts were made to simplify these models and make them more practicable.

The first meaningful simplification of the portfolio theory was the single-factor model proposed by William Sharpe (1963). It was based on the assumption that returns on security are linearly related to some underlying market index. It means that when a market index goes up, most stocks tend to increase in price. Similarly, when the market goes down, the prices of most stocks tend to fall. The return of each instrument depends on the performance of that underlying index, which can be written as follows:

$$R_i = \alpha_i + \beta_i R_m + \varepsilon_i;$$

where:

R_i – The rate of return for asset i,
α_i – The constant term for asset i,
β_i – The slope coefficient for asset i,
R_m – The rate of return for the aggregate market index,
ε_i – The random variable, $\mathrm{E}(\varepsilon_i) = 0$.

As shown in the equation, William Sharpe considerably reduced the number of operations needed to calculate an instrument's return. Since the expected value of the random component (ε_i) is 0, the instrument's rate of return depends essentially on the parameters alpha (α_i) and beta (β_i). Alfa is an independent measure which does not relate to market movements. The beta parameter, in turn, is a measure of the sensitivity of a stock to market movements. It indicates how the return on an individual stock will change when there will be a change in the overall market return. The model proposed by Sharpe can also be illustrated graphically as a regression line called the security characteristic line (see Figure 3):

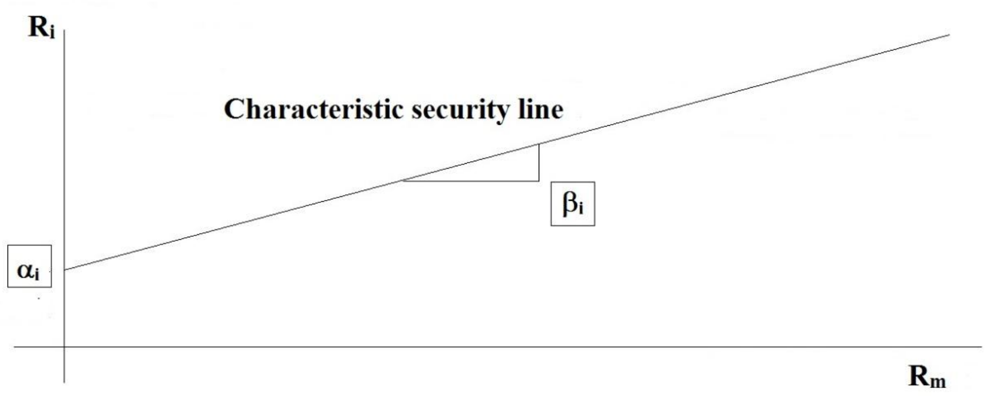

Figure 3. Characteristic security line
(Source: Sharpe, 1963)

The beta parameter determines by how many units the return on security will increase (or decrease) when the return on the underlying index increases (or decreases) by one unit. It might be calculated as follows:

$$\beta_i = \frac{cov_{i,m}}{\sigma_m^2};$$

where:

β_i – The i-th stock beta parameter,

$cov_{i,m}$ – Covariance of the i-th stock with the market,

σ_m^2 – Variance of the market returns.

The beta parameter can be interpreted in the following way:

- When β = 1 the security behaves in the same way as the market. A 1% increase (or decrease) in the market return results in the same 1% increase (or decrease) in the security's return.
- When β > 1, the security responds to the market changes more strongly than the broad market. This means that a 1% increase (or decrease) in the return of an index will result in a greater than 1% increase (or decrease) in the return of the security.
- When 0 < β < 1, the security responds to the market changes less strongly than the broad market. This means that a 1% increase (or decrease) in the return of an index will result in a lower than 1% increase (or decrease) in the return of the security.
- When β = 0 the security does not react to market changes in any way. This means that an increase (or decrease) in the market return does not lead to any change in the return of the security.
- When β < 0, the security reacts to the market changes in the opposite way to the broad market. This means that when the return on the market index rises, the return on that security falls, and when the return on the market falls, the return on the security rises.

The one-factor model shows that the security total risk consists of systematic and unsystematic risk. The systematic risk depends on the risk of the overall market and cannot be reduced. The non-systematic risk is associated with a security-specific risk and could be eliminated by adequate portfolio diversification. A total security risk might be calculated as follows:

$$\sigma_i^2 = \beta_i^2 \sigma_m^2 + \sigma_\varepsilon^2;$$

where:

$\sigma_i^2 -$ variance of the return on the i-th security,

$\sigma_\varepsilon^2 -$ variance of the rate of return on the random component.

From the perspective of passive portfolio management, the key implication deriving from the single-index model is that the market portfolio represents a completely diversified portfolio. It means that in the market portfolio, unsystematic risk is fully eliminated through diversification and consequently, the portfolio consists only of systematic risk, defined as the variability of all the risky assets caused by macroeconomic factors (Reilly and Brown, 2012: 212).

Another crucial extension of the portfolio theory was invented by Sharpe (1964), Lintner (1965) and Mossin (1966) independently. The theory known as the Capital Asset Pricing Model (CAPM) allows investors to determine the expected rate of return for any risky asset in an efficient market. The classical form of the CAPM model describes the behaviour of a capital market in equilibrium and is based on a set of assumptions, including those in Modern Portfolio Theory (MPT) and some additional ones:

1. All investors behave rationally in the sense of the Markowitz model, i.e., they choose a portfolio of assets located on the efficient frontier and tangent to their utility curve.
2. Investors can borrow or lend any amount of money at the risk-free rate of return (RFR).
3. All investors have the same expectations, and they estimate identical probability distributions for the future rates of return.
4. All investors have the same investment horizon.
5. All financial assets are perfectly divisible, i.e. one can acquire any quantity of them.
6. There are no transactional costs or taxes.
7. There is no inflation or any change in interest rates, or inflation is fully anticipated.
8. Capital markets are in equilibrium.

Due to the fact that many of the above assumptions are hardly possible to apply in practice or are even unrealistic, the model is highly questionable. However, in spite of the CAPM limitations, it remains the most popular model used to explain the expected returns of risky assets.

In the CAPM model, two relationships are of primary importance: the capital market line (CML) and the security market line (SML). In a market equilibrium, rational investors invest in efficient portfolios, i.e., portfolios on the capital market line (CML) which is presented in Figure 4. The CML explains the expected rate of return in terms of the price for providing financial capital (RFR) and the price for bearing the total risk of an investment (σ).

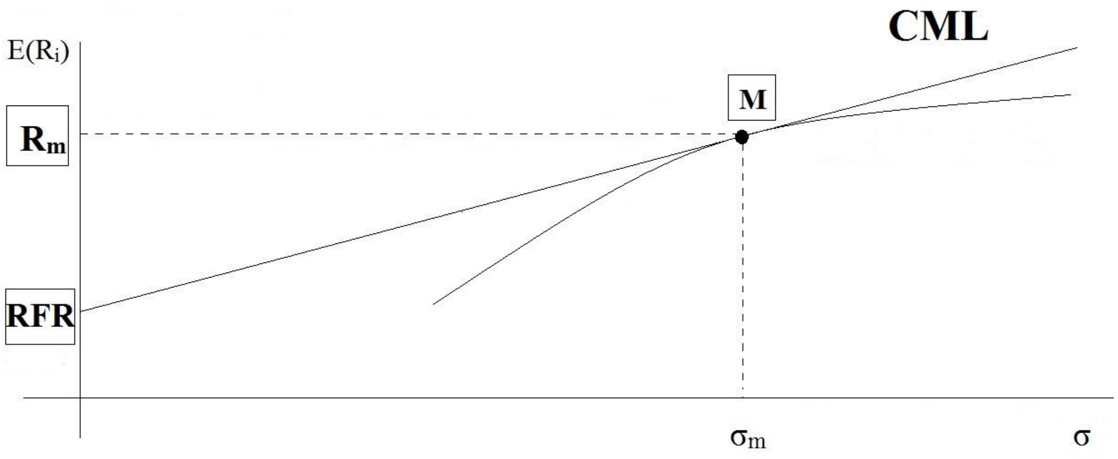

Figure 4. Capital market line
(Source: Sharpe, 1964:437)

A risk-free asset with no variance (RFR) represents an investment whose return is entirely certain. Moving right on the capital market line means increasing the share of risky assets in the portfolio. Portfolio M stands for a market portfolio which is fully diversified and thus its total and systematic risks are equal. The CML might be described by the following expression (Reilly and Brown, 2012: 218):

$$E(R_P) = RFR + \sigma_p \left[\frac{E(R_m) - RFR}{\sigma_m}\right]$$

Another key relationship in the CAPM exists between the beta coefficient and the portfolio return and is shown by the security market line (SML) (see Figure 5). Since investors are assumed to hold efficient portfolios, they attempt to diversify the portfolio to reduce specific risk. Therefore, it can be assumed that the specific risk tends to zero and thus the total risk depends only on the specific risk measured by the beta coefficient (β_i). The implication is that an investor seeking an optimal investment should consider the relationship between the expected return of a portfolio and its beta coefficient.

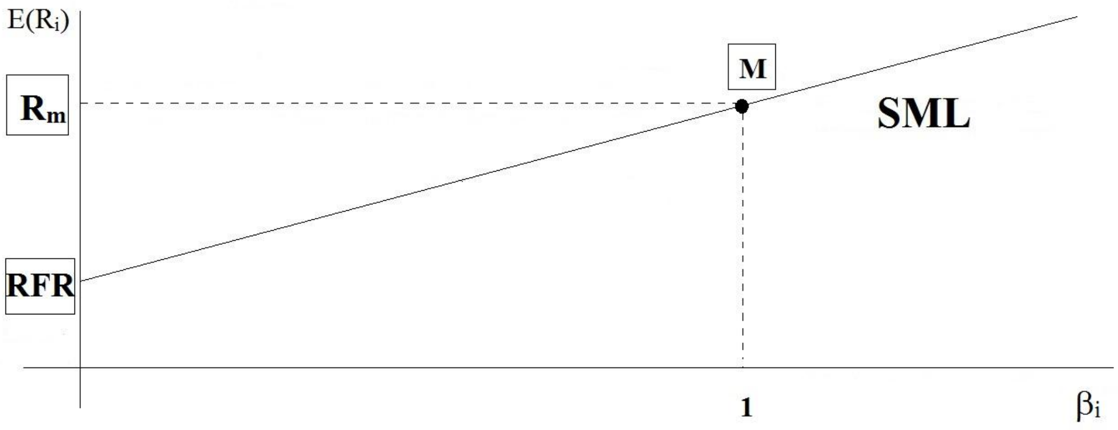

Figure 5. Security market line
(Source: Reilly and Brown, 2012: 218)

The CAPM redefines risk in terms of the security beta coefficient which expresses the non-diversifiable portion of the stock risk in relation to the whole market. The Capital Asset Pricing Model describing a market in equilibrium could be presented by the equation:

$$E(R_i) = RFR + \beta_i[E(R_m) - RFR]$$

The above expression known also as a security market line (SML) differs from the CML not only in terms of the risk measure but also the portfolios considered. The capital market line refers only to efficient portfolios, while the SML considers all portfolios, including those consisting of individual securities. What is common for the mentioned lines is that both the CML and the SML describe the expected rate of return as a sum of the risk-free rate and the expected risk premium.

According to the Capital Asset Pricing Model (CAPM), in a market equilibrium, all capital assets are properly priced. This means that, from the perspective of risk-return relationships, no individual stock is more attractive than another. The model implies that the most rational investment decision is to allocate funds to the market portfolio, as any less diversified portfolio would be suboptimal. This provides the theoretical foundation for passive portfolio management, understood as a strategy that aims to replicate the return of the market portfolio, which is a broad market index.

Roll (1977) criticised the testability of the CAPM, and hence the rationale of the model itself, by showing that for any efficient portfolio, the equation of the CAPM model is an identity. Roll's criticism shows that it is impossible to test the CAPM model, but at most to verify that the portfolio used to derive the β parameters is efficient. Conceptually, the CAPM model appears correct, however, in reality, it is only able to correctly value capital assets if we define the real market portfolio, which in practice turns out to be extremely difficult.

The response to the objections to the CAPM was the Arbitrage Pricing Theory (APT) developed by Ross (1976) and Roll and Ross (1984). The fundamental postulate of this theory (the law of one price) states that two assets that are the same cannot be sold at different prices. Rational investor behaviour reduces arbitrage opportunities to a minimum, which implies that higher risk-free returns are not possible and ensures equilibrium in the capital market.

This model differs from CAPM as it is less restrictive in its assumptions and assumes that returns in capital markets are shaped by some k-factor process as a linear relationship between asset returns and systematic risk factors. The APT model does not specifically indicate the number of these factors and does not name them, but most often refers to macroeconomic factors or systematic risks that cannot be diversified. These factors may pertain to interest rates, inflation, GDP growth rates, industry-specific variables, or other common factors that affect a broad range of assets. What is key, however, is to measure those factors despite they may remain unidentified, i.e. to determine the sensitivity of returns to these factors. Furthermore, it should be emphasized that the APT model does not exclude the influence of the market factor in shaping returns but states that many other factors may also be relevant.

The Arbitrage Pricing Theory (APT) does not specifically justify passive investing as a preferred investment strategy, nevertheless some aspects of the APT might be considered as the rationale behind passive portfolio management. Firstly, the market efficiency in the model deriving from the limitations of the arbitrage mechanism is many aspects is aligned with the presumed market efficiency in the passive investing approach. Then, similarly to passive investing, the APT model advocates diversification across numerous risk factors, and investing in the broad market portfolio might be perceived as one method to do so.

Despite numerous studies questioning the CAPM or its assumptions, the conclusions deriving from the model were of great importance for the finance theory, particularly the passive portfolio concept aiming exactly at tracking the return and risk of the market portfolio. In contrast, the APT model has not been as widely adopted and accepted, primarily due to practical application difficulties. According to Rosenberg et al. (1985), the assumptions of the CAPM met with much criticism since it only relies on market risk to measure stock returns. Therefore, over the years, academics have introduced consecutive extensions to the original CAPM model. Eugene Fama and Kenneth R. French (1993) extended the CAPM into a three-factor model, adding two essential factors: the size factor (SMB) and the value factor (high-minus-low, HML). Later, Mark Carhart (1997) expanded this into a four-factor model by incorporating a momentum factor. Subsequently, Fama and French (2015) introduced a five-factor model, adding profitability and investment factors to their earlier framework. Finally, Fama and French (2018) extended the five-factor model by reintroducing the momentum factor, resulting in a six-factor model (Fama and French, 2018). However, all of these extensions ultimately stem from the original, single-factor CAPM model.

1. 2. 3 Efficient Market Hypothesis

Having considered the essentials of portfolio theory and capital market models, it is time to discuss the most influential theory from the perspective of passive portfolio management: the efficient market hypothesis. Concerning financial investment theory, market efficiency is understood as informational efficiency. In an efficient capital market, all the new information regarding a security is immediately fully reflected in its price, and, as a consequence, the prices of securities reflect their genuine value. Aside from informational efficiency, one may distinguish operational and allocative efficiency. Operational efficiency depicts the ability of a market to facilitate the execution of trades quickly and at low costs, while allocative efficiency refers to the ability of a market to allocate resources to their most productive uses. However, since informational efficiency is fundamental for capital markets, the following considerations will focus on this dimension of efficiency.

Even though the efficient market theory was developed in the 1960s, its origins can be found much earlier. The efficient markets were explicitly mentioned even in 1889 in George Gibson's book *The Stock Markets of London, Paris and New York.* He stated that when “shares become publicly known in an open market, the value which they acquire may be regarded as the judgment of the best intelligence concerning them” (Gibson, 1889). In 1900, Louis Bachelier published his PhD thesis *Theorie de la Spéculation* (Bachelier, 1900) in which he described the assumption of random price formation in the capital market and indicated that “the mathematical expectation of the speculator is zero”. The implications of the efficient market hypothesis were also remarked by the English economist John Maynard Keynes (1923) who said that investors in the financial markets are rewarded not for knowing better than the market, but rather for taking risks.

A number of studies pointed to the randomness of stock prices. Alfred Cowles (1933) analysed the performance of investment specialists and concluded that stock market forecasters are unable to forecast prices because they cannot beat the market. Holbrook Working (1934) published an article *A random-difference series for use in the analysis of time series,* in which he compared the stock returns to numbers from a lottery. Maurice Kendall (1953) analysed 22 price series at weekly intervals and, to his surprise, found that they were essentially random. He concluded that, since individual stocks behave differently from the average of similar stocks, it is not possible to predict stock market movements a week in advance without external information.

Arnold B. Larson (1960) used a new method for those days of time series analysis and noted that the distribution of price changes is "very nearly normally distributed for the central 80 per cent of the data, but there is an excessive number of extreme values". Jay M. Berger and Benoit Mandelbrot (1963) proposed a new model for clustering errors in telephone circuits that can be applied to stock trading and provide a justification for the Pareto-Levy distribution of stock price changes. Paul Cootner (1964) edited his classic book entitled *The Random Character of Stock Market Prices,* comprising a collection of the papers previously mentioned and others.

For the first time the term efficient market was explained by Eugene Fama in his article *The behaviour of stock market prices* (1965). Not only did Fama establish the foundations of the efficient market concept, but he also provided research results indicating that stock prices follow a random walk. In the same year, Paul Samuelson (1965) gave the first formal economic argument for efficient markets in his article *Proof that properly anticipated prices fluctuate randomly*. The term "efficient markets hypothesis" was introduced by Harry Roberts (1967), who was also the first to propose a distinction between weak and strong tests of market efficiency.

The pivotal work on the efficient market hypothesis proved to be Eugene F. Fama's article entitled *Efficient capital markets: A review of theory and empirical work* (Fama, 1970). Fama provided a synthetic summary of the previous theoretical and empirical achievements related to efficient market models and formulated a theory of efficient markets. As far as the majority of the prior papers explained the market efficiency in terms of the random walk, Fama's theory was based on a fair game model. He defined an efficient market as "a market in which prices always fully reflect available information". Therefore, the emergence of new information on the market does not contribute to achieving excessive returns for investors. Importantly, for his ground-breaking research on efficient markets, in 2013 Fama was awarded the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel.

The efficient market hypothesis (EMH) is based on a set of assumptions. First of all, it assumes a large number of participants who analyze and value securities independently from the others in order to maximise their profits. Secondly, the inflow of new information is random, which implies that investors are unable to predict future information. Thirdly, the buy and sell decisions of all the investors lead to a rapid adjustment in security prices to reflect the effect of the new information. Finally, most importantly, the expected returns on securities fully reflect their risks. Consequently, the current price of a security is consistent with its risk-return characteristics.

Fama (1970) distinguished three forms of the efficient market hypothesis depending on the scope of information included: the weak-form EMH, the semistrong-form EMH and the strong-form EMH. The weak-form EMH relates only to the relationship between security market information and the price of a stock. It assumes the current price fully reflects all the market-related information about security including the past returns, trading volume, sequence of the prices and any other relevant market data. One consequence of the weak-form efficient market hypothesis is that the future prices might not be predicted on the base of the historical prices. Therefore, any trading strategies and technical analysis methods do not ensure successful investment results.

The semistrong-form EMH extends the scope of information disclosed in the price of a security to all public information. It assumes that every publicly announced information on a security is immediately reflected in its price. Not only does it include the market information, but also the company financial reports, earnings, and dividend announcements, news about the economy, political changes, and many others. As a result, the hypothesis implies that neither technical nor fundamental analysis guarantees above-average risk-adjusted returns.

The strong-form EMH assumes that the security price reflects all the public and private information. The hypothesis includes the assumptions from weak-form EMH and semistrong-form EMH and expands it to the insider company data. It contends that markets are perfectly efficient, which means that all information is cost-free and publicly available to everyone at the same time. Consequently, every investor has access to a full range of relevant data concerning securities. The acceptance of this hypothesis would also mean that any group, including insiders, could not beat the market. It implies that none of the active portfolio management strategies could bring abnormal returns and thus rationalizes the passive portfolio management.

Having explained the fundamental assumptions of the efficient market hypothesis and its subhypotheses, it is time to discuss some results of EMH tests. Beginning from the weak-form EMH, it is typically checked by two groups of methods: statistical tests of independence and tests of trading rules. It was Fama (1965) who used run tests to support the random walk model and demonstrated that stock prices change over time independently. Brealey et al. (2011)

conducted a study on a set of blue-chip companies where the correlation coefficient of return on two consecutive days ranged from -0.03 to 0.03, and they concluded that the historical stock return does not influence the following returns. However, Allen et al. (2006) analysed weekly returns and showed the lack of dependencies.

The second group of weak-form EMH tests relies on comparing the risk-return results derived from a simulation of a particular trading rule of technical analysis and a *simple buy-and-hold* strategy. Fama and Blume (1966) conducted tests on the effectiveness of various filters on the *buy-and-hold* policy for Dow Jones Industrial Average's stocks and inferred that filters do not contribute to excessive returns. Parks and Zivot (2006) analyzed both the returns of individual companies and equity indices. They stated that the technical analysis could not be profitable due to the existence of transaction costs.

Coming to the semistrong-form EMH, it is commonly tested by studies to predict future returns on the base of public information and event studies examining how fast the significant economic events are reflected in the security price. Both tests consider the abnormal rates of return, which are adjusted by the market return to extract the sole return on security without the market's impact on it (Reilly and Brown, 2012: 155). Ball and Brown (1968) investigated 261 companies for the effect of annual earnings announcements by using the time-series regression method and evidenced that only 10–15% of the information has been anticipated. On the contrary, the results achieved by Jones, Rendleman and Latané (1982) revealed that 51% of the total response in stock returns came after the earnings announcement. One study concerning the stock split information showed that such an event is reflected in security price even before the public announcement of it so investors could not achieve above-average returns by following publicly revealed news (Fama et al., 1969). Alford and Guffey (1996) analysed the seasonality in 18 countries over the periods 1970–1994 and 1983–1994. They concluded that seasonalities may not exist which implies that investors cannot take advantage of utilizing the patterns to predict future price movements.

The primary objection to semistrong-form EMH pertains to the existence of market anomalies which are discrepancies between the expected market behaviour under the assumption of market efficiency and the real market behaviour (Shiller, 1980). Market inefficiencies include seasonal effects, overreaction and mean-reversion (DeBondt and Thaler, 1985). Furthermore, value and momentum effects are frequently reported as a piece of evidence for the semi-strong market efficiency rejection (Asness, Moskowitz and Pedersen, 2013).

The strong-form EMH tests include the analysis of the investment performance of corporate insiders, security analysts and professional money managers (Reilly and Brown, 2012: 166). The hypothesis has been indirectly supported by a great number of studies comparing the results of actively managed mutual funds to market portfolio performance (see section 1.3.1 for more details). However, the vast majority of direct tests on the strong-form EMH did not confirm it. Lorie and Niederhoffer (1968) studied transactions over the period 1950–1960 and concluded that when insiders accumulate stocks, their rate of return can be expected to outperform the broad market within the next 6 months. Jaffe (1974) reported that insiders' returns exceeded the performance of non-insiders by approximately 2.5% in the years 1962–1968. Studies by Chowdhury, Howe and Lin (1993) and Pettit and Venkatesh (1995) also showed that corporate insiders consistently were able to achieve the above-average returns by possessing private information. On the other hand, Meulbroek (1992) stated that the stock market detects and discounts transactions related to illegal insider trading before the information on such a transaction becomes public. There were also studies concerning the professional analysts' performance, which contradict the EMH. For instance, Womack (1996) indicated that analysts have special timing and stock-picking abilities, apparently as a result of acquiring the internal news.

The Efficient Market Hypothesis remains a subject of interest to researchers from around the world to this day. According to Robert J. Shiller, laureate of the 2013 Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel, the EMH is a "half-true" since it excellently describes trading conditions in the modern stock market but fails to explain the existence of some patterns in stock price behaviour. One important argument against the EMH is based on the participants' irrationality, which is also the subject of behavioural finance. According to Coval and Shumway (2005), it is the loss aversion that contributes most to the participants' unreasonable decisions. Baker and Nofsinger (2010) mentioned loss aversion, overconfidence, anchoring, framing, and confirmation bias as the most common examples of investor irrational behaviour. Malkiel, Mullainathan, and Stangle (2005) indicated that asset bubbles are the result of investors' failure to maintain a short position in a mispriced asset. Another contradiction to the EMH refers to investors' overreaction to unexpected corporate news. It means that stock prices tend to fall more than they should by the information value. Battalio and Mendenhall (2005) noticed that investors tended to underestimate financial announcements. On the contrary, Mishkin and Eakins (2012) indicated the market overreaction to unforeseen unfavourable news.

The concept of an efficient market, as disputable as it is, constitutes one of the most significant theories in modern finance. Despite numerous contradictions, including market anomalies, excess stock volatility, and overreaction, as well as the occurrence of asset bubbles and investors' irrational decisions, the EMH could not be fully rejected. The principal arguments for the acceptance of the efficient market hypothesis derive from the fact that stock returns are commonly found to be random, and investors are not able to achieve above-average returns on their investment in the long run. All in all, it is the strong-form EMH that laid the foundations for passive investing and contributed to the development of passively managed exchange-traded funds.

1. 3 Active vs. Passive Asset Management

The term asset management encompasses a wide range of practices aiming to maximise the value of investment while maintaining an acceptable level of risk. It is divided into two principal schools of thought: active management and passive management, which will be discussed in the following part of the chapter.

From a historical standpoint, active investing was the first known approach to portfolio management. Active asset management presumes to outperform the benchmark on a risk-adjusted basis. Active managers' goal is to produce a positive alpha (α), which constitutes an added value that he or she add to the investment process. This approach to investing assumes beating the market by constructing a portfolio in which actual returns surpass risk-adjusted expected returns. An additional value to the active portfolio might derive from tactical adjustments, such as specific investment strategy and market timing or security selection, which includes picking the winning stocks. The active manager collects information, uses forecasting techniques, updates the real value of assets, and makes decisions regularly. His or her abilities are the predominant conditionality for the above-average risk-adjusted returns (Brentani, 2004: 93).

Passive investing, which has its roots in efficient market hypothesis, assumes investing in the entire market. Under the theorem that markets efficiently allocate capital to different companies, passive portfolio managers focus on the replication of the weights within the market in their own portfolio. Passive investors accept that one cannot beat the market, so their goal is to achieve the market average return, which they accomplish by mimicking the performance of the broad market. As the approach is based on holding in the portfolio assets that track the return of the benchmark index over time, it is also called indexing.

Typically, passive investing is associated with a long-term buy-and-hold strategy with occasional rebalancing in case of changes in the underlying benchmark composition. The strategy does not attempt to predict short-term market movements, but instead emphasizes the importance of staying in the market over extended periods. This gives patient investors the opportunity to benefit from the long-term growth and compounding of their investments. As in the long run, the economy grows, an investor locating money in the broad market would also gain in the longer time horizon.

Importantly, the purpose of passive portfolio management is to generate the rate of return and risk that best match these characteristics from the tracked benchmark. Consequently, a passive manager is judged by his or her ability to minimise the deviation in return and risk between the constructed portfolio and the index portfolio. If the manager attempts to outperform the benchmark results, he or she breaches the premise of a passive portfolio (Reilly and Brown, 2012: 550).

There are also modern instruments that represent a combination of active and passive portfolio management, such as smart-beta funds. Building on factor-based extensions of the CAPM model, they aim to achieve return enhancement, risk reduction, and increased diversification compared to traditional market-capitalization-weighted index investing across systematic risk factors, like value, momentum, size, and low volatility. Thus, despite utilizing an index as a benchmark, they are considered active instruments. FTSE Russell's (2017) report indicated that smart-beta funds constitute a less expensive alternative to traditional active investing.

Finally, it should be emphasized that passive investing might be understood slightly differently in the academic discourse and practice. Essentially, scientists identify a passive strategy with investing in a broad capitalization-weighted market portfolio. This means that any strategy that breaches the premise of investing in the broad market in accordance with the weightings in the market portfolio is considered active. Consequently, investing, e.g., in a specific market sector via instruments that mirror the performance of a sector index might not be perceived as a passive strategy. On the other hand, practitioners commonly define passively managed sectoral index funds or ETFs as passive investment vehicles.

1. 3. 1 Key Research on Active vs. Passive Investing

The debate on the rationale for passive versus active investing has been continuing for years. The first papers exploring this issue were published in the 1960s by three well-known researchers: Sharpe, Treynor, and Jensen. Treynor (1965) analysed 20 U.S. mutual funds in the years 1953–1962 and indicated that, on average, their results were poorer than the Dow Jones Industrial Average (DJIA) index performance. Sharpe (1966) conducted a study on 34

U.S. mutual funds operating in the period 1954–1963 and found that only 11 funds outperformed the DJIA index while 23 funds turned out to be inferior. Jensen (1968) investigated 115 U.S. mutual funds in the period 1945–1964 and concluded that the average rates of return of the funds were lower than the returns of random buy-and-hold portfolios after the risk consideration.

Several studies on actively managed mutual funds conducted in the subsequent years confirmed active funds' inferiority in comparison to index portfolios (Chang and Lewellen, 1984; Henriksson, 1984). In the 1990s, however, several authors began to suggest that the earlier research was biased by a survivorship effect, i.e., they included only funds that remained in the market throughout the entire period of study. Brown, Goetzmann, Ibbotson, and Ross (1992) pointed out that studies with samples that contain only funds that survived over the research period could apparently display performance persistence even though none actually existed. According to Barras et al. (2010), there were systematic differences in the data on mutual funds' performance before 1996.

Nevertheless, having improved the research method to eliminate the survivorship effect, the great majority of studies carried out on the American market still demonstrated the superiority of passive funds. A study by Malkiel (2003) showed that from 1991 to 2001 around 70% of the mutual funds received lower returns than their benchmarks. Several years later, after increasing the sample, Malkiel (2011) evidenced that 66% of the U.S. mutual funds in the period 1970–2010 had lower returns than their benchmarks. He also noted that mutual funds that were most profitable in the short term did not maintain their superiority in the long term. Moreover, mutual fund managers who earned the highest returns in one year achieved only an average return in the next year. Brealey et al. (2011) found that the U.S. mutual funds were able to achieve higher returns than their benchmarks only in 16 of the 47 years studied.

Similar studies have also been conducted in other developed markets, as well as in emerging markets. Blake (1998) thoroughly researched the performance of 2300 UK open-ended mutual funds over 23 years and demonstrated the underperformance on a risk-adjusted basis by the average fund manager. Kaserer and Pfau (1993) and Scherer (1994) examined German equity mutual funds and concluded that these funds underperform the market. Otten and Schweitzer (2002) compared the performance of European and U.S. equity mutual funds and identified the relatively poor performance of U.S. funds in comparison to European funds and the superior performance of small-cap funds. Christensen (2005) analysed the performance of 47 Danish funds between January 1996 to June 2003 and reported that fund managers had neither special selection nor timing skills.

There were also contradictory research results over the years, however, it should be noted that the empirical evidence for the superior performance of active investing is not as strong as the evidence for the advantage of passive investing. Carlson (1970) recalculated Sharpe and Jensen's results using annual fund returns from 1948 to 1967 and obtained contrasting results. Furthermore, he pointed out that the conclusions derived from fund returns depend largely on the period of research, the type of fund, and the benchmark chosen. Another evidence of actively managed funds' superiority includes a study by Kon and Jen (1978) who showed that mutual funds could achieve abnormal results. One controversial paper includes research by Ippolito (1989), who investigated 143 mutual funds in the years 1965–1984 and stated that risk-adjusted returns on mutual funds (excluding acquisition fees) were larger than the index funds' performance. However, the results were not confirmed by Elton, Gruber, Das and Hlavka (1993), who examined all equity mutual funds that existed for the period 1965–1984 and evidenced that on average these funds underperform index funds. As Lehman and Modest (1987) noted the principal reason for different outcomes in mutual funds performance pertains to the benchmark choice.

Several studies provided empirical evidence on the increased performance of active funds during economic recessions. Moskowitz (2000) estimated that over the period 1975–1994, the average annualized return achieved by mutual fund managers was 1% higher in recessions than in non-recessions. Fortin and Michelson (2002) noted that index funds outperformed actively managed funds on a total return and risk-adjusted basis. However, in the periods of going into and out of recession, active funds tended to achieve higher returns. Kosowski (2011) found that the mutual fund risk-adjusted returns during recessions exceeded those in expansions on average by 3–5% per year. Staal (2006) indicated that in the years 1962–2002, the average fund's risk-adjusted return was negatively correlated with the Chicago Fed National Activity Index. Finally, Glode (2011) concluded that U.S active equity mutual funds generated remarkably better performance when the economic conditions were poorer.

Some books have played a significant role in spreading the passive investing approach. The first seminal contribution was a work by Malkiel (1973) titled *A random walk down Wall Street*. The author focused on the analysis of the random walk hypothesis with regard to capital market investments. He argued that the acceptance of this hypothesis in practice entails the impossibility of predicting short-term share price movements. This, in turn, implies that analysing complex price formations, forecasting company earnings and even using investment advisory services is pointless. According to Malkiel, neither technical nor fundamental analysis enables investors to generate higher rates of return than a simple *buy-and-hold* strategy. In a later edition of the book from 2003, Malkiel proved that during the 25 years analysed, more than two-thirds of active funds underperformed the passive Standard & Poors's 500- Stock Index Fund. Furthermore, he remarked that the performance of active fund managers is extremely variable and argued that the probability that a manager who beat the index in one year would maintain his or her position in the subsequent year is less than 50%. As a result of his research, Malkiel concluded that it is the long-term *buy-and-hold* strategy which tracks the performance of a broad stock index, that is the most profitable investment approach. He argued that due to the high management and distribution costs of mutual funds, it was likely to be the passive *buy-and-hold* strategy that would yield the highest rate of return. All in all, Malkiel transferred the idea of passive investing from the academic world to the practice of fund management and influenced the thinking of many industry leaders who pioneered index funds, including John C. Bogle.

Another remarkable publication that contributed to passive investing prevalence was a book written by Charles D. Ellis *Investment policy. How to win the loser's game* (1985). The author notes that the presumption that institutional investors can beat the market is fundamentally false, as they are the ones who make the market and cannot overcome themselves. Ellis indicates that the transactions made on the New York Stock Exchange by mutual funds in the 1960s represented only about 10% of all transactions, while in the 1970s they already comprised approximately 90%. Market changes made institutional investors no longer compete with individual investors, but rather with each other. The enormous increase in management costs induced by growing competition has forced managers to achieve higher and higher rates of return to cover these costs. Paradoxically, increasing efforts by fund managers to outperform the broad market have made this even more challenging. Indeed, most managers underperform the market in the long term when costs and commissions are taken into account (over the 25 years studied, more than 75% of professional funds underperformed the S&P 500 index). By and large, Ellis (1998: 6) advises that since active fund managers are unable to beat the market over the long term, they should consider joining the market by investing in index funds that replicate the performance of the broad market.

Last but not least book supporting the passive approach to investing is *The Little Book of Common Sense Investing* by John C. Bogle (2007). Bogle was a founder of the Vanguard Group which nowadays is one of the largest investment companies in the world and a very pioneer of index funds. His principal intention was to make investing easier and less expensive. According to Bogle, the most reasonable investment strategy is to adopt a *buy-and-hold* policy which is based on purchasing low-cost index funds that track the broad market index such as

S&P 500. The book comprises a polemical argument in favour of index funds and their superior position as compared to active mutual funds. Furthermore, the Author highlighted the importance of minimising costs and stated that “costs make the difference between investment success and investment failure” (Bogle, 2007: 45). By and large, John Bogle as a highly influential professional significantly contributed to the passive investing industry proliferation and the popularisation of passive funds among individual investors.

Research on the profitability of active and passive funds is conducted not solely by academics but also by industry professionals. One of the most recognized publications investigating the returns of actively managed funds in comparison to indices is the SPIVA U.S. Scorecard. It has been prepared annually and continuously since 2002 by S&P Dow Jones Indices. Not only does the survey encompass domestic equity fund performance, but also the results of international equity funds and fixed-income funds. What is more, the study distinguishes additional subcategories. Domestic equity funds are divided into large-cap, mid-cap, and small-cap funds depending on the size of companies the fund invests in, and value, growth, and core funds concerning the fund's investment strategy. Among the international funds, in turn, a distinction is made between global funds, international funds, international small-cap funds, and emerging market funds. The principal measure used in the report is the percentage of funds underperforming the benchmark. The study provides detailed data on the funds' performance over annual, 3-year, 5-year, 10-year, and 20-year periods.

According to the SPIVA Scorecard report (2022), 95.39% of U.S. domestic equity funds underperformed the S&P Composite 1500 index over the past 20 years. Moreover, when the risk-adjusted returns are taken into account, this figure is equal to 95.39%. When it comes to the S&P 500 index and the large-cap equity U.S. funds performance, the report showed that over the 5 years, 59.22% of funds underperformed the benchmark, and over the last 20 years, 94.12% of funds. On a risk-adjusted basis, 95.65% of actively managed large-cap equity funds achieved inferior results over the 20 years as compared to their benchmark. Similarly, the performance of global funds turned out to be poorer than the index S&P Global 1200. Over the 5 years, 69.17% of global funds underperformed the benchmark, and on a risk-adjusted basis, 74.17% of funds. Consideration of the 20 years showed that 87.70% of global funds achieved a lower return per unit of risk than the index (see Table 1).

Table 1. Percentage of U.S. equity funds underperforming benchmarks

Fund category	**Comparison index**	**3-year (%)**	**5-year (%)**	**10-year (%)**	**20-year (%)**
All Domestic Funds	S&P Composite 1500	80.33	80.85	92.68	95.39
All Large-Cap Funds	S&P 500	72.53	74.80	90.20	95.65
All Mid-Cap Funds	S&P MidCap 400	46.73	54.69	70.25	90.10
All Small-Cap Funds	S&P SmallCap 600	53.72	62.54	80.33	93.61
Large-Cap Growth Funds	S&P 500 Growth	88.31	85.83	98.05	99.71
Large-Cap Value Funds	S&P 500 Value	64.92	64.24	84.74	80.09
Mid-Cap Growth Funds	S&P MidCap 400 Growth	26.52	33.09	60.47	93.03
Mid-Cap Value Funds	S&P MidCap 400 Value	54.10	69.23	78.65	77.55
Small-Cap Core Funds	S&P SmallCap 600	62.92	75.76	88.41	92.23
Multi-Cap Value Funds	S&P Composite 1500 Value	87.60	86.89	89.71	85.12

Note: Based on risk-adjusted returns, data as of Dec. 31, 2021. Source: S&P Dow Jones Indices LLC

1. 3. 2 Significance of Costs

Even assuming the possibility of active funds to generate abnormal returns due to market inefficiencies, after the consideration of costs, active funds commonly offer investors lower returns. Costs are essential in the calculation of an investor's net returns since they reduce the overall profits. It is especially important in long-term wealth accumulation as even minor differences in costs would compound over the years.

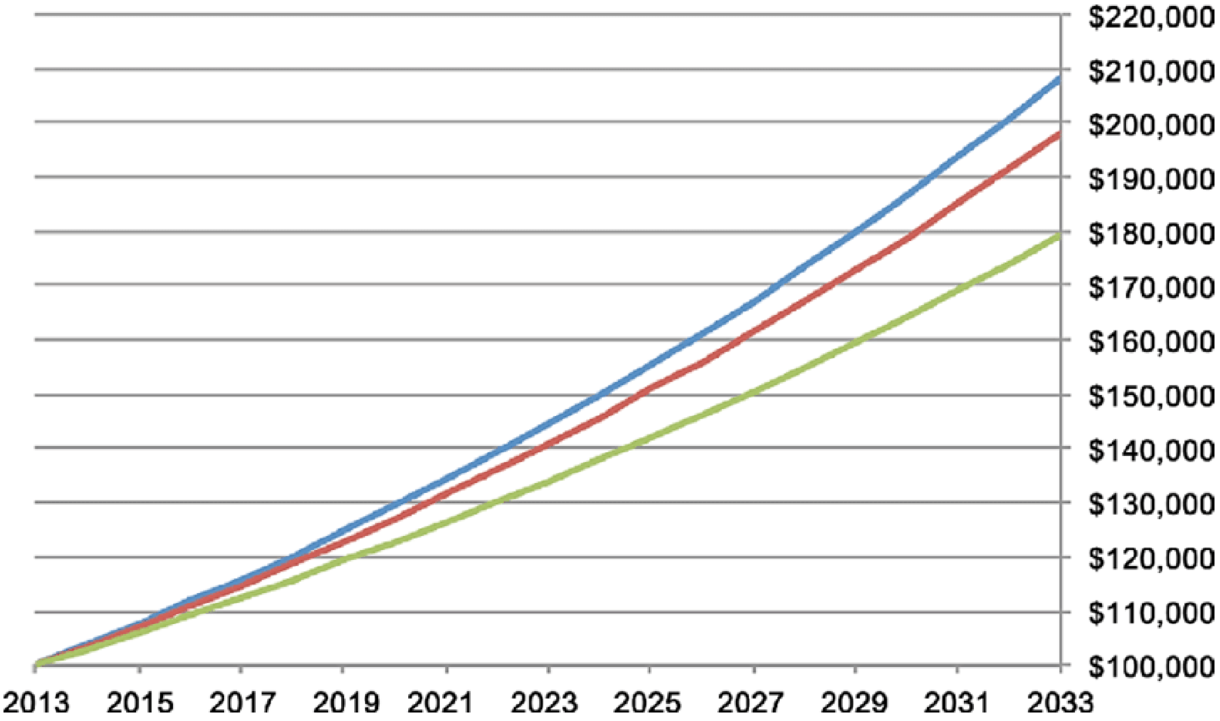

Figure 6. Portfolio value from investing USD 100,000 over 20 years
(Source: Securities and Exchange Commission 2014)

Note: Blue line = 4% annual return less 0.25% annual fees, red line = 4% annual return less 0.50% annual fees, green line = 4% annual return less 1.00% annual fees.

The Securities and Exchange Commission (2014) demonstrated that even minor differences in fees can compound over time, significantly impacting portfolio returns. Over 20 years, the difference in portfolio value between the lowest-cost option (green line) and the highest-cost option (blue line) amounted to nearly USD 30,000, based on an initial investment of USD 100,000.

It is commonly known that active portfolio management is characterized by considerably higher costs as compared to passive portfolio management (see Figure 7). Possessing new information, constant analysis, active selection of the instruments, as well as frequent shifts in the portfolio holdings, incurs substantial costs. High costs of active fund management are, in turn, reflected in the poorer fund performance and higher fees charged to clients of such funds. According to Sharpe (1991), these are the costs that will always make active investing an inferior alternative. He proves that the average performance of actively managed funds must be lower than that of passively managed funds by an amount equal to the average active fee. Then, Carhart (1997) examined over 1800 mutual funds in the years 1962–1993 and revealed that only the first decile of all actively managed funds earned high enough returns to cover management fees and costs. Shukla (2004) proved that active funds can generate abnormal gross returns. However, due to the costs, they do not outperform passive indices when net returns are considered.

As passive funds seek to generate, on average, a return equal to the market return before costs, the average return of all active investors must be equal to the market return. In contrast, active investors strive to beat the market, so for one investor to beat the market, others must lose. Taking into account additional excessive costs associated with active investing, on average, active investors receive a lower return than passive investors.

One comprehensive study on the costs related to active and passive investing was conducted by French (2008). He compared fees, expenses, and trading costs involved in active investing on the U.S. exchanges (NYSE, Amex, and NASDAQ) between 1980 and 2006 with the estimated cost of investing in an instrument that would mimic the index. The study proved that active investing required spending 0.67% of the aggregated value of the market each year. French concluded that under reasonable assumptions, a shift toward passive investing would increase the annual return of the typical investor by 67 basis points.

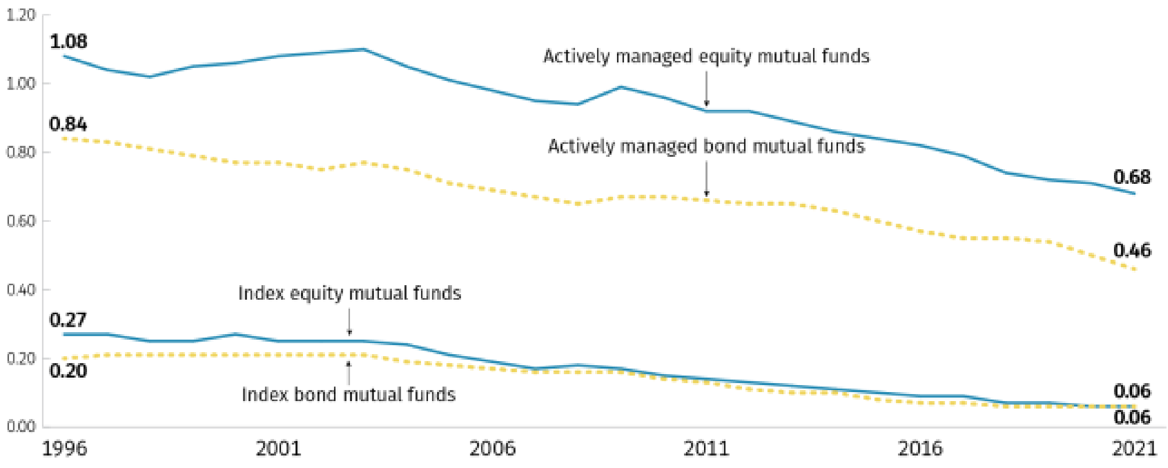

Figure 7. Average Expense Ratios of Actively Managed and Index Mutual Funds in the U.S. Equity and Bond Markets, 1996–2021

(Source: Investment Company Institute, Lipper, and Morningstar, 2021)

What one can observe in recent years is the decline in the cost of both passively and actively managed funds. However, passive funds remain leaders in expense ratio minimization. Undisputedly, this is the core of the rise of interest in passive investing, so it would be useful to investigate the phenomenon.

Passive funds can offer lower expense ratios for their clients for several reasons. Firstly, as they aim to replicate the performance of a specific benchmark rather than actively selecting securities, the management costs are substantially lower. Passive fund managers do not need to engage in extensive research, analysis, and stock picking. Without the need for in-depth research, teams, and analysis tools, index-based funds can significantly reduce their expenses and offer their clients lower expense ratios.

Then, passive funds benefit from reduced portfolio turnover. Actively managed funds are supposed to search for market opportunities related to market inefficiencies, which might occur many times a day or even an hour. This implies that they are forced to incur high brokerage commissions, bid-ask spreads, and market impact costs. Conversely, passive funds are just worried about tracking the index's composition. Depending on an index methodology, rebalancing occurs monthly, quarterly, or even annually, but quarterly is the most common. The differences in frequency of adjusting the portfolio composition and weighting between active and passive funds drive the level of costs incurred by funds and, in turn, determine the expense ratios.

Finally, passive funds managed to reduce their expense ratios over the years as a result of economies of scale. With an increase in assets under management (AUM), their fixed costs, such as administrative expenses, regulatory compliance, or index licensing costs, could be spread across a larger investor base, so funds might offer lower fees for clients. The competitive pressure among passive fund providers was also crucial. Since all index-based funds provide similar products for their clients, costs are the principal reason for a fund to stand out.

Interestingly, the decreasing costs of passive funds seem to be beneficial for the whole mutual fund industry. Cremers et al. (2015) found that competitive pressure from low-cost indexed funds made actively managed funds charge lower fees. Additionally, the pressure made active managers generate higher alpha. As index funds charge significantly lower fees than active mutual funds, more and more investors are prompted to passive investing. In that way, active mutual funds are forced to be more effective to stay competitive.

1. 3. 3 The Rise and Controversies of Passive Investing

It was academic research indicating an underperformance of active funds and relatively higher costs of active investing that made an important contribution to the active-passive shift. The proliferation of the efficient market hypothesis in the 1960s led more and more investors to question the rationale behind actively selecting securities for a portfolio and trying to "beat the market" (Bhattacharya and Galpin, 2011). The introduction of the first passive index fund in the 1970s, and then, from the 1990s, the gradual expansion of passive ETFs, made the passive investment approach practicable for retail investors. Finally, progressive regulation of fees in investment products may have encouraged the financial industry to offer low-cost passive products (Sushko and Turner, 2018).

While in 1995 passive funds represented only 3% of the U.S. market, in 2005 they already accounted for 14%. However, the greatest growth in the passive fund industry was seen after the financial crisis of 2007–2008, mainly due to the proliferation of low-cost exchange-traded funds (ETFs). In 2013, passively managed mutual funds and ETFs already comprised about 25% of the market, and by 2019 their share exceeded 40% (see Figure 8).

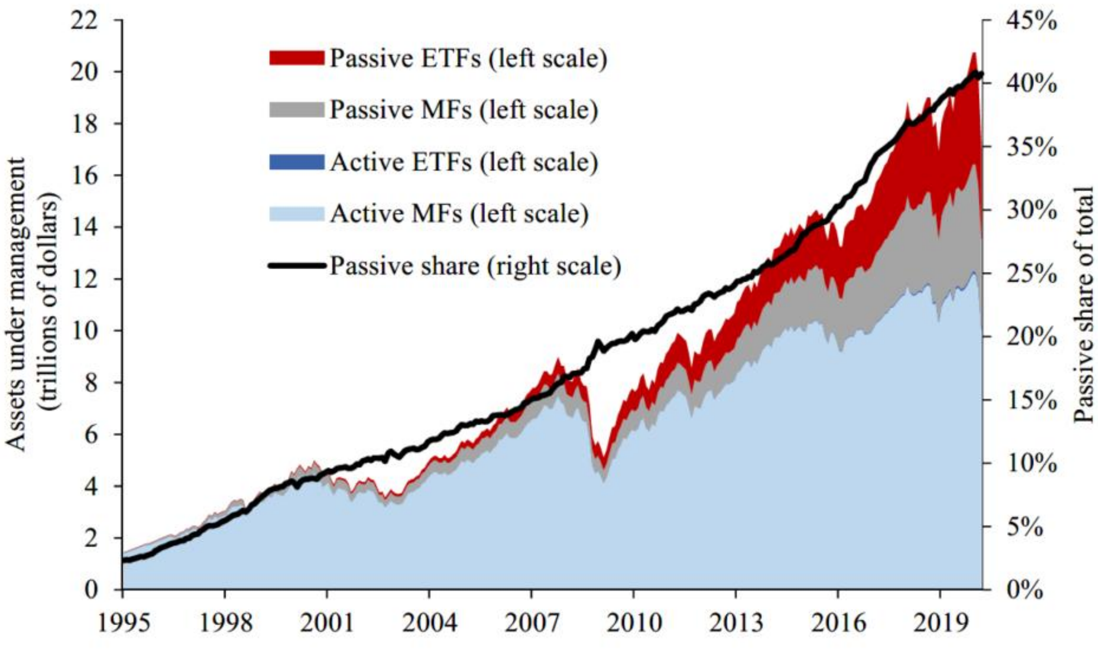

Figure 8. Total Assets Under Management in U.S. Passive and Active Mutual Funds and ETFs, and Passive Share of the Market (1995–2019)

(Source: Morningstar, Inc.)

However, it should be highlighted that there are also many opponents of a passive portfolio management approach. The first objection to passive investing is reduced market efficiency. As passive investing does not involve independent data analysis, research, and investment decision-making, a passive investment approach may lead to less accurate valuations and a general decline in market competitiveness. On the other hand, the Vanguard research *Setting the record straight: Truths about indexing* (Rowley et al., 2018), did not confirm that indexing harms capital markets. They found no evidence that indexing contributes to fewer market opportunities for active managers, causes greater market volatility, or leads to a declining number of public companies.

Fichtner et al. (2017) pointed to the risks associated with the continued growth of ETFs and other passive index funds, which may create new financial risks, including an increase in the herd effect among investors and greater volatility during periods of severe financial volatility. The authors showed that the passive index fund industry is dominated by three giants: BlackRock, Vanguard, and State Street (SSGA). Combined, they are the largest shareholder in at least 40% of all US-listed companies and 88% of the companies in the S&P 500 index. Next, Bebchuk and Hirst (2022) reported that at the end of 2021, the Big Three collectively held a median stake of 21.9% in S&P 500 companies, which represented 24.9% of the votes cast at the annual meetings of those companies.

The implication is that they have considerable room for corporate governance abuse. BlackRock, Vanguard, and State Street may be able to influence the board of directors through private engagement in the pursuit of some of the common goals of large passive index managers. There is even evidence that the voting patterns of the three funds are converging. Bubb and Catan (2022) applied machine learning techniques to analyse a comprehensive dataset of mutual funds' votes and confirmed that BlackRock, Vanguard, and SSGA are all members of the same party which they called the "Traditional Governance Party". Braun (2015) argued that passive investors can act as "patient" capital and thus facilitate long-term strategies. As such, the Big Three have the potential to bring about significant changes in the political economy of the United States by influencing important corporate themes such as short-term versus long-term goals, executive compensation inadequacy and mergers and acquisitions (Fichtner et al., 2017). Another notable criticism of passive investment funds is that they compete primarily on fees and tracking error size and have no incentive to contribute to the value of the companies they hold in their portfolios. Bebchuk and Hirst (2022) raised the voice that competition with other index funds gives index fund managers no incentive to invest in stewardship for any of the companies in their portfolios.

On the whole, the consistently positive capital flows towards passive investments testify to the growing popularity of this form of investment among investors. It is noteworthy that even certain risks associated primarily with overconcentration in the industry do not seem to have deterred investors from continuing to invest passively.

1. 4 Indices as Benchmarks in Passive Portfolio Management

Even though theoretically the market portfolio should contain all the investable assets in the given market, due to practical reasons it does not. Primarily, not all the assets are available for investors to buy. As a result of regulatory restrictions, high minimum investment requirements or limited market liquidity, some assets are not easily accessible. What is really important in the market portfolio is to be a good representation of the broad market so to constitute a good enough representation it does not need to encompass all the assets. In fact, the diversification benefits resulting from including in the portfolio some additional assets would be minor and the increased efforts and costs to do so would be significant. Thus, the market portfolio is considered an approximation in the context of the Capital Asset Pricing Model (CAPM) because it is difficult or almost impossible to construct a portfolio that contains all investable assets in proportion to their market value.

1. 4. 1 Index Functions

From the perspective of the major subject of the paper, indices constitute a base for passive investment products. Essentially, the market portfolio in passive investing is understood as a capitalization-weighted broad market portfolio. This means that the weight of each security is determined by its market capitalization relative to the total market capitalization of all the securities in the portfolio. Securities with higher capitalization are more important and thus represent a broader part of the portfolio.

Nevertheless, one should be aware that financial indices do not emerge as a response to the need for benchmarks for passive products. Originally, financial market indices served only as synthetic indicators of the general conditions prevailing in a given market or its segment. They reflect the average value of a basket of financial instruments and provide all investors, even non-professionals, with an accurate overview of the market trends. It was this informational demand that prompted the development of financial indices, with the first financial index designed in 1884 by the journalist Charles H. Dow, who, together with Edward Jones, operated the firm Dow Jones & Company (Lo, 2016: 1–3).

The pivotal function of market indices is to provide synthetic information about the current market conditions. Indices aggregate a great number of market members with homogenous characteristics in one average measure (Schyra, 2013: 13). In fluctuating market conditions, they enable simple, cost-effective identification of current trends and their comparison nationally and internationally. In other words, financial indices provide a significant gauge of the market participants' attitudes and behaviour. Regularly, indices reflect investors' confidence and market sentiment. It is an essential source of information for all investors and financial market analysts required to evaluate the market situation and make investment decisions. Furthermore, market indices are frequently used by economists and academics as a barometer of future business activity. They allow researchers to compare the historic values of the securities and thus investigate the long-term market performance and correlations between different asset classes. The results of the analysis are of key importance for effective portfolio management (Ferri, 2008: 94). Taking into account the constantly growing number and complexity of financial instruments in the contemporary economy, it appears that today the informational function of indices does not lose its significance and remains their basic application.

With the development of financial markets, indices began to perform new functions. One important function of indices is benchmarking – providing a reference point against which the performance of the investment portfolio can be measured. This is especially pertinent to the evaluation of the portfolio management efficiency and the measurement of the mutual investment institution's performance. At this point, it is necessary to mention that even though the terms index and benchmark are commonly used synonymously, in fact, they are not the same. An index is a hypothetical investment portfolio pertaining to the performance of the entire market or its segment, whereas a benchmark is a standard point of reference – a portfolio representing an acceptable rate of return for a given level of risk (FTSE Russel, 2017: 1). A financial market index can be used as a benchmark to assess the abilities of an active portfolio manager compared to the performance of the broad market (Schyra, 2013: 12). Not only is this relevant for clients of mutual funds, pension plans or insurance funds, but also for the boards of such institutions to monitor the performance of their employees.

Then, in the context of the subject matter of this paper, a key function of indices is to serve as an instrument for the construction of index financial products. By utilising financial market indices, convenient and transparent financial instruments are created for the low-cost implementation of a passive investment strategy. For this reason, financial indices are licensed by fund managers to be used as the basis for passive products that track index performance and risk. As a result, the index is not just a theoretical portfolio reflecting the market's condition, but a real product enabling investment in the index. The first index funds were created in the

1970s, and the passive exchange-traded fund (ETF) industry gradually expanded from the 1990s. Today, by purchasing a single investment product, an investor can easily benefit from exposure to the global equity market. The development of passive investment products has been enabled by financial indices, while over time the constantly growing interest and demand for passively managed products have boosted the index industry even further (Schyra, 2013: 15). What is more, there are other instruments traded on today's financial markets that are linked to indices. This includes ETN products (exchange-traded notes), ETC products (exchange-traded commodities), as well as some derivatives like index options and tracker certificates.

The problem with today’s indices is that more and more often, they lose their primary function associated with being a market approximation. As far as the target market, an index is obliged to follow was selected based on the asset type, and geographical region, nowadays it commonly uses other characteristics like company size or investment style. This means that indices have started to apply the active management component. More sophisticated indices serve as a base for more inventive passive products. However, more complicated does not mean better. For instance, fundamental indexation is an alternative approach to constructing indices, where the weights of the securities are based not on market capitalization (as in traditional indices), but on fundamental metrics such as earnings, book value, revenue, cash flow, or dividends (Arnott et al., 2005). Estrada (2006) argued that fundamental indexation is not necessarily the best approach to obtaining international diversification. From the theoretical perspective, dividing the market into pieces is not justified as long as it does not represent an independent asset class or a market with unique risk-return characteristics. However, as a response to the growing demand for more imaginative passive products, index providers broaden their product range.

1. 4. 2 Benchmark Index Construction

Having discussed what financial indices are, what their purpose is, and how they work, it is time to consider the characteristics of the index that underpins passive investment instruments. Benchmark indices constitute stable baskets of securities and are constructed to measure the performance and risk of a given market or its segment. It is a simple representation of the market investment opportunities with weightings compatible with the real share of the market for each constituent security (Ferri, 2008: 88).

It was Sharpe (1992) who first indicated what criteria should be met by an index to evaluate the performance of an actively managed portfolio properly. He noted that a benchmark index ought to constitute a viable, inexpensive, and identifiable alternative to an established active portfolio. What is more, it should not be easily outperformed by another portfolio.

Comprehensive guidelines for benchmark construction were provided by the CFA Institute – a global non-profit organisation associating professional investors from over 100 countries. To ensure relevance, integrity and consistency of benchmarks they have formulated the following benchmark index characteristics (Ferri, 2008: 85–87):

1. Simple and objective selection criteria – the security inclusion rules should be unambiguous and clear.
2. Comprehensive – the index should encompass all the realistically available securities.
3. Replicable – the index must represent a realistic portfolio which could be followed by a passive investor.
4. Stability – the composition of the index should not be altered often, and all the changes need to be predictable.
5. Relevance – the index should track those markets or market segments which are most interesting for investors.
6. Barriers to entry – the markets or market segments tracked by the index should not be characterised with significant entry barriers.
7. Expenses – to measure the market performance adequately, the benchmark index should not involve excessive or unpredictable expenses.

Siegel (2003) indicated that the most crucial benchmark index characteristics is market-capitalization weighting. Furthermore, he lists the following features of a good benchmark index:

- completeness – the broader and deeper the market coverage of an index, the more complete it is;
- investability – the index should consist only of securities which can be effectively purchased by investors;
- clear, published rules and open governance structure – to ensure the predictability of benchmark index moves as a result of the particular market events;
- accurate and complete data – investors should be able to review key data concerning the benchmark index;
- acceptance by investors – the benchmark index should be accepted by a broad range of investors;

— availability of crossing opportunities, derivatives, and other tradable products – the benchmark index should ensure accessibility to a variety of investment products offered by institutional providers;

— low turnover and related transaction costs – to facilitate the tracking ability of the entities replicating the benchmark index.

However, in practice, complying with all the above-mentioned conditions simultaneously proves to be extremely difficult or even impossible, as some of them are contradictory. Financial market indices commonly encounter a bias problem (Andreu, 2009). Firstly, the sample bias may lead to the formation of market indices which characterized by different return and risk characteristics even though they aim to represent the same analytical market (Amenc and Martellini, 2003). Then, the selection bias depicts heterogeneous inclusion and exclusion criteria for securities to be an index component. Finally, the float adjustment methodology might affect the representativeness of an index. Because information about the availability of a security in the free float is not always public, divergences between indices representing the same market are likely to occur. The reason for this includes numerous trade-offs that index providers deal with.

Siegel (2003: 8–10) pointed out several trade-offs in the benchmark index construction:

- Completeness vs. Investability. Theoretically, the ideal index should include all securities listed on the tracked market or its segment. However, investing in every single security in a given market, disregarding its liquidity and accessibility, would affect the investability of an index. Thus, replication of such an index would prove to be excessively costly and inefficient.

- Reconstitution Frequency vs. Turnover. Because an index reconstitution entails turnover, it is a source of costs to investors. Timely reconstitution ensures adequate tracking of an asset class by an index but also involves higher trading costs. As a result, high turnover implies a poorer ability of passive funds to replicate the index performance.

- Rebalancing Frequency vs. Turnover. Like reconstitution, rebalancing generates trading costs and all its consequences. This is why index providers determine the rules on how often rebalancing is carried out, e.g., quarterly or annually.

- Objective and Transparent Rules vs. Judgment. The advantage of the indices with strictly defined objective rules includes the predictability of the changes in the index composition and weighting, and thus efficiency in the index replication process. On the other hand, judgement-based indices offer greater flexibility and enable the index to be given specific desired features.

Another significant consideration in index construction is weighting. The weight given to each member of the sample determines how much of the constituent security to include in the index. Thus, the weighting method used by an index provider significantly affects the value of an index. According to Ferri (2008: 127), one might differentiate three major security weighting methods: capitalization weighting, fundamental weighting, and fixed weighting (see Table 2).

The capitalization weighting assumes that the market knows best what importance each constituent security should have. One important advantage of the capitalization weighting method is that it maintains the proportions from the target index value. This approach ensures that larger companies have a greater impact on the index performance, aligning with their relative importance in the market. Capitalization-weighted indices inherently incorporate the market's assessment of the relative value and significance of different securities and thus do not alter the real valuation of companies. Having in mind the Modern Portfolio Theory and the consistency with the long-term buy-and-hold strategy, capitalization weighting is the only justified weighting method (Andreu, 2009). Additionally, the method is indicated to have low turnover costs (Schoenfeld, 2004; Arnott et al, 2005), which justifies its application in the construction of a benchmark for passively managed funds.

On the other hand, the principal disadvantage of the method is associated with overweighting large companies, constituting the dominant part of the market. Potentially, this may lead to overexposure to specific sectors or industries. What is more, in the short term, capitalization-weighted indices tend to overweight stocks in which prices have risen and underweight stocks in which prices have fallen recently.

Table 2. Classification of security weighting methods

Capitalization	**Fundamental**	**Fixed Weight**
Full cap	Dividends	Equal
Free float	Financial	Modified equal
Constrained	Price	Leveraged beta
Liquidity	Momentum	Short beta
Production	Qualitative	Long/Short

Source: Portfolio Solutions, LLC.

Fundamental weighting methods aim to determine the weight of each constituent security in an index by using the factors from the corporate financial statements. These factors include such fundamental measures of company size as book value, dividends, earnings, number of employees, and many others. Hsu and Campollo (2006) claimed that fundamental indices deliver significantly higher returns than their respective capitalization-weighted benchmarks. However, Ferri (2008) argued that due to their complexity, they are not supposed to serve as a benchmark portfolio for the performance of active strategies or as a measure of market returns. He stated that fundamental-weighted indices have only one relevant use: they are commonly used as the basis of investment products. Finally, according to Arya and Kaplan (2007), any non-capitalization weighted index often requires rebalancing and thus induces excessive transaction costs.

There are also fixed weighting methods that assign equal weight to every constituent security in the index portfolio. Undoubtedly, fixed weighting methods are characterised by great simplicity and thus are easy to calculate. Moreover, studies by DeMiguel et al. (2009) and Plyakha et al. (2015) proved that equal-weighted portfolios outperform other weighting strategies. However, opponents of this approach enumerate its numerous disadvantages, including overweighting the securities that represent a small fraction of the target market and underweighting the securities that represent the largest fraction of the target index. What is more, equal-weighted indices demand frequent adjustments to the index since any price change of the constituent security makes the index no longer equally weighted (CFA Institute, 2020: 287). Interestingly, according to Malladi and Fabozzi (2007), the excess returns of equal-weighting are higher than the higher costs resulting from higher portfolio turnover, which shows that equal-weighting makes economic sense.

All in all, it might be stated that there is no one universally accepted framework for constructing a benchmark index. Even though there are essential rules for appropriate index construction, the rules applied by benchmark index providers depend on the underlying function to be performed by the index. This is why the index industry offers a wide range of products that aim to align with various investment strategies.

1. 4. 3 Index Industry

The entities responsible for constructing and managing financial market indices (index providers) are not a homogeneous group. Conversely, it is a highly diverse category of institutions, comprising both companies that operate stock exchanges (like Intercontinental

Exchange) and specialised companies that focus only on providing indices (like S&P Dow Jones Indices (SPDJI), MSCI, FTSE Russell).

The number of financial market indices is growing rapidly. According to the Index Industry Association (2022), the number of indices worldwide exceeded 3 million. To better understand the magnitude of financial market indices, one might explore their division (see Table 3). It gives clear evidence of the diversity and increasing sophistication of market indices today. One index is described by all the categories simultaneously. This means that differences between two indices just within one category make them unique products, even though all the other characteristics remain common.

Table 3. Division of financial market indices by type

Criteria	**Asset class**	**Market segment**	**Geography**	**Market capitalization**	**Investment style/ strategy**
Types	equity	broad market	regional	large-cap	market-cap
	fixed income		country-specific		
	commodity		global		
	real estate	sectoral		mid-cap	equal-weighed
	interest rate				price-weighted
	currency				factor-based
	derivatives	industry	emerging markets	small-cap	dividend-weighted
	multi asset		developed markets		smart beta
	alternatives		exchange-based		ESG/ Sustainable

Source: Own elaboration based on major index providers' websites.

Ferri (2008: 110), who was a founder of Portfolio Solutions LLC, provided classification and examples of security selection categories and methods (see Table 4). He divided the index security selection methods into three broad categories: passive, screened, and quantitative. Passive security selection methods pertain to a simple replication of a market or market segment after eliminating illiquid securities that are not relevant for an index performance. They are used by all benchmark index providers, for example, in the global index construction. Next, the security screening methods are based on applying special filters to screen a large universe of securities. The filters are limited only by the index providers' imaginations and might include any security characteristics, from the fundamental qualities, through specific industries or regions, to the elimination of unethical companies. Then, quantitative security selection methods intend to beat the broad market by choosing the securities with the highest probability of superior performance. These methods commonly use sophisticated models and artificial intelligence techniques to achieve abnormal risk-adjusted returns (Ferri, 2008: 110–124).

Table 4. Classification of security selection approaches in index construction

Passive	Screened	Quantitative
Full replication	Exchange	Economic Cycle
Sampling	Fundamentals	Multifactor
Buy and hold	Price Trend	Momentum
Single Issue	Thematic	Proprietary*

*Proprietary means not enough public information is available to determine

Source: Portfolio Solutions, LLC.

Not only is the index industry growth visible in the increasing number and types of indices, but also in industry revenues. According to Burton Tailor International Consulting (2022), the index industry revenue grew 23.1% in 2021 and amounted to USD 5 billion. Equity indices remain the major source of index providers' revenue, with a 64.6% share of index providers' total revenues. The main industry drivers include ESG investing, factor investing, as well as crypto indexing (Burton Tailor International Consulting, 2022: 7–10).

Another essential feature of today's global index industry is a high level of concentration. Both in Europe and the United States, four major index providers for exchange-traded funds (MSCI, SPDJI, FTSE Russell, and Bloomberg) account for more than two-thirds of the market shares. While in Europe the dominant index provider is MSCI, in the United States it is SPDJI. The concentration is of great importance for the ETF industry as ETF sponsors rely on index providers. According to An et al. (2022), 60% of index licensing fees paid by U.S. equity ETFs are markups. The authors evaluated that the elimination of the index providers' market power could reduce ETF management fees by 30%.

1. 5 Summary

The issues presented in this chapter related to the theoretical foundations of passive portfolio management, the specifics of passive investments and their rationale, as well as indices as a basis for constructing passive portfolios, are essential to the topic of this dissertation. Without the development of capital market theory, subsequent research on the efficiency of investment funds, as well as the growth of the index industry, the development of passive ETFs would not have been possible.

The reviewed literature indicates that, from the perspective of finance theory, the most efficient investment portfolio is a market portfolio whose best approximation is a broad equity market index. Moreover, the acceptance of the efficient market hypothesis entails that low-cost passively managed investment portfolios replicating the performance of a broad market index are the most viable investment approach for most investors over the long term. The implication from the vast majority of mentioned studies on the efficiency of active and passive funds is that the costs associated with acquiring market information and actively managing a portfolio are not worth it, which justifies the rationale for adopting a passive strategy.

Then, the chapter presented the essence of financial market indices. Importantly, index functions are not limited to the benchmark for passively managed investment products, like passive ETFs, but also serve as a proxy for active portfolio managers and inform about the current state of the economy. Moreover, the chapter outlined the magnitude of market indices and the methods of their construction. When it comes to index providers, they represent an enormously diverse group, however, the industry is dominated by four major players (MSCI, SPDJI, FTSE Russell, and Bloomberg). This has negative consequences for costs incurred by entities licensing indices, primarily financial institutions managing passive funds.

Although the theory presented in this part of the dissertation does not fully comprehend the broad and complex problem of passive portfolio management and is merely an outline of the most relevant issues, it is sufficient to understand the theoretical basis of the thesis. It provides an important starting point for the considerations on ETFs, as well as the problem of the tracking ability of passive equity ESG ETFs.

CHAPTER 2 ESG INVESTING

2. 1 Introduction

Although ESG investing is a relatively new concept, it has already started to play a large role in the investing industry. The term "ESG" was first used in 2004 by the United Nations Environmental Program Finance Initiative in its publication *The Materiality of Social, Environmental, and Corporate Governance Issues to Equity Pricing*. By the beginning of 2020, more than 97% of funds invested on a sustainable basis were allocated to ESG strategies (US SIF Foundation, 2020). Over the 15 years of its existence, the ESG investing strategy has already changed considerably. While initially it was concentrated on ESG screening, which mostly deteriorated the performance, over the years, the industry introduced ESG ratings, and now, materiality metrics.

Several factors have contributed to the rapid growth of ESG investing. First and foremost is climate change and governmental and international actions to counter it. The Paris Agreement (2015) is a legally binding international treaty on climate change adopted by 196 Parties at the 2015 United Nations Climate Change Conference (COP 21) in Paris. The principal goal of the agreement is to limit global warming to well below 2°C, preferably to 1.5°C, compared to pre-industrial levels. It is intended to reach the global peaking of greenhouse gas emissions as soon as possible to achieve a climate-neutral world by 2050.

Political influences are increasingly relevant in the proliferation of the ESG concept. Geopolitical tensions, populism, and trade wars can effectively affect corporate behaviour. Finally, sanctions imposed on individual corporations, industries, or states harming the environment and society enforce compliance with certain standards. Public pressure on legislators and governments is also of great importance in popularising ESG investments. As a result of numerous protests and consumer boycotts, unethical actions can be stopped. One example is the behaviour of states and customers after the Russian attack on Ukraine. The deliberate avoidance of purchasing products and services from Russian-owned companies effectively led to the loss of competitive advantage or even bankruptcy of many companies linked to activities incompatible with the ESG framework.

Another significant determinant of the development of ESG investments has been the progress in ESG measurement and reporting standards. Initiatives undertaken by the United Nations or the Sustainability Accounting Standard Board substantially contribute to corporate awareness and facilitate transparent reporting of environmental, social, and governance issues.

Furthermore, the advancement in data and technology has allowed more comprehensive, detailed, and flexible ESG data analysis.

The following chapter deals with ESG investing considerations. Firstly, it places the ESG strategy among other ESG-related approaches and strategies. Secondly, it explores the three dimensions of the ESG acronym. Then, it focuses on the major ESG investing and reporting initiatives and the key research on ESG investment performance and risk. Finally, it investigates the passive ESG approach that essentially constitutes the base for passive ESG ETFs.

2. 2 ESG Terminology

As a relatively new investing approach, ESG investing and terms related to this approach tend to be misunderstood or confuse. Quite often terms describing different approaches to investments seeking positive social and environmental outcomes in addition to financial performance are used interchangeably which causes chaos. Because no strict rules or standards distinguish between various approaches to responsible investing, in practice, different strategies tend to overlap. Nevertheless, to place the ESG investing approach and its strategies among the broader world of responsible investing, the following part of the dissertation will be devoted to classifying the ESG-related terminology.

2.2.1 Defining ESG investing

Even though the term ESG investing is commonly used among financial professionals and academics, it seems that some discrepancies in what ESG investing means occur among different groups of its users. Taking the example of major ESG product providers, we might observe relevant inconsistencies. One Vanguard Research Report, *ESG, SRI and impact investing: A primer for decision-making* (Grim and Berkowitz, 2020) places ESG investing among the broader concept of sustainable investing. ESG investing is defined as an investment-related activity that accounts for some type of environmental, social, or governance consideration. Responsible investing, socially responsible investing (SRI), thematic investing, and sustainable investing are mentioned as some related terms. Such an approach to defining ESG investing leads to mixing various approaches, as there are no strict frames where one approach finishes and another begins. Essentially, sustainable investment focuses on the selection of assets that can be defined as sustainable, i.e., contributing in some way to a long-term vision of sustainability. This is a broad term that encompasses numerous approaches to responsible investing, from ESG integration to impact investment, and actually might mean everything.

The MSCI (2018) publication *Introducing ESG Investing* remarks that ESG investing is often used synonymously with sustainable investing, socially responsible investing, mission-related investing, or screening. Their definition of ESG investing entails the consideration of environmental, social and governance factors alongside financial factors in the investment decision-making process. They notice that behind ESG investing may stand various investor needs: improvement of financial performance, aligning with personal beliefs, or generating a positive impact on the world. This implies that the term ESG investing is treated as a package for multiple strategies that consider both ESG and financial factors. Johnson et al. (2020) pointed out that ESG is an umbrella term that covers many previous concepts from Corporate Social Responsibility (CSR), Sustainability, Environmental, Health and Safety (EHS), to Corporate Social Performance (CSP).

There is no common academic definition of ESG investing, however, the merit of ESG focuses on the integration of non-financial factors into the investment process. What distinguishes ESG investing from other approaches to sustainable investing is the materiality and financial relevance of ESG factors. Principally, in the ESG investing approach, non-financial factors should have a significant impact on a company's financial performance and risk profile. The consideration of the ESG financial implications is intended to ensure long-term sustainability and potential for financial outperformance. Then, as ESG investing relies on material factors, ESG data and ratings are of key importance in this approach. What is more, ESG investing does not assume any separation from conventional financial considerations but rather the expansion of traditional investment analysis. Thus, it is about the integration of environmental, social, and governance factors alongside financial fundamentals.

While conventional investment primarily aims to maximize the financial return for a given level of risk, sustainable investing additionally entails consideration of non-financial factors such as environmental, social, or corporate governance aspects. The particular approaches to sustainable investing demonstrate significant differences in the emphasis on financial and positive social impact issues. While ESG investing prioritizes maximizing the financial outcome, impact investing pays particular attention to the social outcome of the investment, with the financial results placed only in second place. Socially responsible investing values financial outcomes and positive impact equally in search of consensus. In opposition to conventional investing is philanthropic activity, whose single most important goal is to maximize positive social impact (Swedroe and Adams, 2022: 8). As shown in Figure

9, different approaches to sustainable investing might be distinguished based on the principal focus of the investment. ESG investing is closely related to conventional investing, which suggests that the primary aim of ESG investment is to deliver a decent return on investment. The difference is that, additionally, it takes into consideration non-financial factors which, in theory, should bring some positive impact on society and the environment.

Conventional Investing	Sustainable Investing			Philanthropy
	ESG Investing	Socially Responsible Investing	Impact Investing	

Financial Focus ⟷ Impact Focus

Figure 9. Sustainable Investing Spectrum
(Source: Swedroe and Adams, 2022:8)

The current state of ESG investing is the result of a longstanding evolution. Focusing on the development of ESG investing, firstly we should explore the ethical and faith-based investing. Ethical investing allocates capital in accordance with certain values and principles. Most often, this approach is based on negative or exclusionary screening of companies involved in activities deemed by the investor to be unethical or contrary to certain international declarations, conventions and voluntary agreements. These exclusions might include tobacco, alcohol, pornography, weapons or a significant breach of international agreements (Hill, 2020: 273).

Faith-based investing, known also as religious investing relates to an investment activity which is in line with some moral or social beliefs. This approach does not concentrate on delivering high or even market returns, but rather on following some religious imperatives and guidelines. Usually, this is reflected in the avoidance of so-called “sin stocks” or others incompatible with the faith and the support of companies consistent with their beliefs. For example, investing based on Christian values may involve avoiding companies that facilitate abortion, contraceptives, embryonic stem-cell research or the production and sales of weapons. Furthermore, investors seeking investments that comply with the Christian religion may often favour companies that care about human rights, employee welfare, or environmental responsibility. Meanwhile, investing based on Islamic values may involve avoiding investment in companies that carry heavy debt loans, profit mainly from interest, or in pork-related businesses (CFA Society of the UK, 2021: 10–11).

Accordingly, for many years, socially responsible investing (SRI) was the dominant strategy for responsible investing. This approach, in addition to financial factors, considers the wider social and environmental context and considers the long-term social impact of the investment. Like a conventional investment, SRI seeks to maximise return for a given level of risk. However, it additionally extends the analysis to include non-financial factors such as environmental, social, and corporate governance aspects. According to Swedroe and Adams (2022: 8-9), this approach might be considered balanced as it seeks both a return for the investor and the opportunity to incorporate into the investment process values consistent with responsible investment. Initially, socially responsible investing was limited to the negative screening of companies that were not in line with the investor's values. Disinvestment commonly included companies in the tobacco, alcohol, weapons, gambling, or polluting industries. Consequently, it often led to excessive exclusions, reduced portfolio diversification, and lowered the total return of the portfolio. Over time, SRI has expanded to incorporate positive screening and has begun to actively invest in companies that implement pro-social and pro-environmental solutions (Hill, 2020: 14). One S&P Global (2020) publication *What is the difference between ESG investing and socially responsible investing?* explains that while ESG investing explores the ESG risks and opportunities that can have material impacts on companies' performance, SRI investing takes financial returns in the second place after the investors' moral values have been considered.

Then, green investment is an approach to responsible investing in which the investor consciously selects "green assets". This includes funds, equities, bonds or infrastructure projects that in some way contribute to solving environmental problems and facilitate the transition to a low-carbon economy. The purpose of green investment is to allocate capital in assets in which the underlying business mitigates problems associated with climate change, biodiversity decline, inefficient resource management and other environmental challenges. Examples of such environmentally friendly technologies and processes are alternative energy, low carbon power generation and vehicles, smart grids, energy efficiency, pollution control, recycling, or waste management (CFA Society of the UK, 2021: 9).

Green investment is a very broad term encompassing the various subcategories of thematic investment and impact investment. To determine the level of "greenness" of an investment, one may utilise the Shades of Green methodology provided by the Center for International Climate Research (CICERO). The methodology reflects the extent to which investments and operations contribute to the transition to a low-carbon and climate-resilient future and their exposure to climate risks. The Shades of Green range from Dark Green to Red and encompass the following (CICERO Shades of Green, 2021: 8):

— Dark green – is assigned to projects that are fully aligned with the long-term vision of a net-zero and climate-resilient future. These initiatives typically involve transformative, low-emission technologies and demonstrate strong environmental governance. An example includes large-scale wind energy developments that are embedded in robust sustainability frameworks.

— Medium green – represents investments that make substantial progress toward climate objectives but do not fully meet the criteria of long-term alignment. These may include, for example, certified green buildings that demonstrate high energy efficiency but may still rely on conventional technologies or supply chains.

— Light green – covers activities that are generally climate-friendly and contribute to emissions reductions but fall short of transformative change. These projects are often considered transitional in nature. An illustrative case is the cleaner or more energy-efficient production of emissions-intensive materials, such as cement or steel.

— Yellow – designates projects that offer limited or ambiguous climate benefits. These activities do not actively support the low-carbon transition and may be associated with ongoing greenhouse gas emissions or exposure to climate risks. An example might include incremental efficiency improvements in fossil fuel infrastructure.

— Red – applied to projects that counteract climate goals, locking in high emissions and long-term climate risk. These include, for instance, investments in new coal-fired power plants or the expansion of coal-related infrastructure.

The range of shades that is applied to the assessment depends on the service. For Company Assessments and Sustainability Linked Bonds, CICERO assigns all five shades, whereas for Green Bonds and Loans and Sustainability Bonds and Loans *are graded Dark Green, Medium Green, or Light Green (CICERO Shades of Green, 2021: 9).

Then, thematic investment refers to investing in assets or products that fall under a specific sustainable-related theme. However, not every thematic investment might be considered a responsible one. Any thematic investment aims to invest capital in companies that are connected to long-term trends in the economy. Whether a thematic fund is a responsible investment depends not only on the theme but also on the ESG performance of the underlying companies (CFA Society of the UK, 2021: 9).

Sustainable-themed investment may pertain to climate change mitigation, clean technology, sustainable agriculture, healthcare, and others. This approach presumes that allocating capital to entire industries, especially relevant to the transformation towards a sustainable economy, is more sensible than selecting individual companies. The motives for thematic investment may include the investor's desire to support the particular industries, as well as the urge to earn high returns on the most forward-looking businesses (Swedroe and Adams, 2022: 23). Although thematic investing may resemble a simple sector investment strategy, the approaches differ significantly. Most notably, thematic investing typically covers multiple sectors and selects companies within those sectors that are relevant to the theme (Investment Leaders Group, 2014: 55).

Another ESG-related category is impact investing, which is a subset of responsible investing that refers to such investments that have a specific positive and measurable environmental and social impact. What distinguishes impact investing from philanthropy is the intention to generate a return for an investor alongside bringing concrete pro-social and pro-environmental changes. It covers a wide range of initiatives undertaken by investors, but generally can be divided into two broad categories (Ormiston et al., 2015):

— "Financial-first" - this approach refers to an investment primarily intended to obtain a financial return that is market-competitive, and additionally results in some social and environmental benefits;

— "Impact-first" – this approach refers to an investment that notably seeks high social and environmental impacts, accepting below-market returns.

Fundamentally, impact investing emphasizes "blended value" by pursuing to achieve both financial returns and measurable social impacts (Bugg-Levine and Emerson, 2011). However, it is a flexible investment strategy that allows for prioritizing one or another. Depending on the impact investors' intentions, their expectations regarding risk, return, and impact vary significantly (Ormiston et al., 2015: 356). As the Global Impact Investing Network's annual survey showed, 66% of investors in impact investing pursue competitive, market-rate returns (GIIN, 2019).

2.2.2 ESG Strategies

Having placed ESG investing among other investment approaches, it is time to discuss various strategies incorporated in ESG investing. Starting from the oldest and easiest strategies to implement, one will revoke negative exclusionary screening. The strategy pertains to excluding some companies, sectors, subsectors, or even countries that are not aligned with ESG criteria. Usually, the criteria are based on ESG ratings, ESG controversy scores, the percentage of revenue associated with a screened issue, or the true and false criteria. The major problem with this strategy is that excessive screening would lead to lowered returns. Thus, the core objective of negative screening is to exclude companies that do not meet specific ethical, social, or environmental criteria, while at the same time preserving a portfolio's overall risk-return characteristics comparable to those of conventional investment strategies (Kumar et al., 2019).

Another way to incorporate an ESG strategy is positive screening. Most often, the strategy pertains to a best-in-class investment, which involves purchasing companies scoring better ESG metrics than their competitors. It is based on selecting only the companies that overcome a defined ESG ranking hurdle within each industry or sector. The foundation of a typical investment using the best-in-class method is to assess companies from different industries from an ESG perspective, assign them appropriate weights depending on the sector, and then construct a portfolio from qualified companies (CFA Society of the UK, 2021: 8). The undoubted advantage of this approach is diversification. As this strategy captures the best performers in terms of ESG factors from various sectors, it provides exposure to a wide range of sectors and is not focused on one industry that has a high ESG score overall (Swedroe and Adams, 2022: 21). Not only can best-in-class investment target specific pro-environmental and pro-social industries, but also more controversial sectors, such as thermal coal, and invest in companies that are leaders in their industry in terms of meeting environmental, social and governance criteria significant for the business.

One new approach connected to positive screening is the ESG momentum strategy, which strives to pursue companies whose ESG scores rise faster than their peers. By overweighting securities whose ratings have increased in recent periods, the strategy aims to generate excess returns (Kumar et al., 2019). This implies that the ESG momentum strategy is short-term oriented as the market is expected to react quickly to ESG score changes (Nagy et al., 2016). What is more, one significant drawback of the strategy is that it does not go for companies with the best ESG ratings but rather with those with the fastest ESG score increase, so the securities might have scores even below the average.

Next approach – ESG integration is a strategy that incorporates environmental, social and governance qualities in an integrated approach along with traditional financial factors. This approach aims to correctly identify, evaluate and price environmental, social and economic risks and opportunities. The underlying rationale behind this approach is the recognition that social, environmental, and governance issues can impact the risk, volatility, and long-term return of an investment. Importantly, ESG integration does to assume overvaluation of the ESG factors importance. Instead, ESG considerations are treated as part of the investment process which do not solely determine the decision (BlackRock, 2022: 5). Ideally, it should consider each level of a company's operation, from its business model via product strategy, distribution system, research and development to its human resources policy. In practice, ESG strategy is implemented through a variety of techniques, and how robust the ESG analysis is depends on the reliability of data, analyst background and experience (CFA Society of the UK, 2021: 5). According to the survey among institutional investors conducted by Eccles et al. (2017), themost severe barriers to the incorporation of the full ESG integration strategy include the lack of standards for measuring ESG performance (pointed out by 60% of respondents) and the lack of ESG performance data reported by companies (pointed by 53% of respondents).

Finally, the shareholder engagement technique differs significantly from all previously discussed strategies toward ESG investing since it is not in any way related to portfolio composition. The essence of this approach is the active ownership of investors to genuinely influence the actions of corporations on ESG issues. It is a technique that actively encourages corporations to behave more responsibly, for example through dialogues with the board of directors, tailored letters, collaborative campaigns or the way the investor votes at shareholder meetings. The method is focused on preserving and enhancing the long-term value of a business by purposeful actions undertaken by the shareholder (CFA Society of the UK, 2021: 274–275). The efficiency of the engagement technique depends largely on the scale of investor ownership in the target company and its perceived market power (Investment Leaders Group, 2014: 56).

2. 3 ESG Dimensions

The term ESG refers to the wide range of environmental, social and governance issues relevant to the long-term sustainable development of the economy, the creation of shareholder value and the mitigation of business risk. To better understand the ESG investing approach, it would be helpful to describe thoroughly the particular dimensions of this strategy. Although there are currently no universal rules standardising E, S and G considerations, and particular activities may overlap, it is possible to present the key issues of the various ESG dimensions on the basis of existing interpretations delivered by international organisations.

2. 3. 1 Environmental Factors

The connections between economics and the environment are evident even at the definitional level. Economics is commonly defined as "the study of the relationship between ends and scarce means which have alternative uses" (Robbins, 1932: 15). Then, environmental sustainability is widely understood as "seeking to meet the needs and aspirations of the present without compromising the ability to meet those of the future" (Brundtland, 1987). The search for efficient ways to manage scarce resources and the trade-off between current costs and future benefits has been a fundamental problem in economics since its inception.

But since the Industrial Revolution in the 18th century, humanity has begun to exploit natural resources at an increasing rate. For decades, the economy and its development were based on environmental resources. Only since the 1970s have scientists begun to sound the alarm that the consequences of economic activity based on the uncontrolled exploitation of the environment could be dire. This refers primarily to the burning of fossil fuels, as well as other forms of pollution and environmental degradation, all of which represent a real threat to the stability of the climate system with potentially catastrophic consequences for humanity and life on the Earth. Successive reports from the Intergovernmental Panel on Climate Change (IPCC) have played a particular role in the study of climate change. It has been marked that all economic activities ought to be undertaken with respecting "planetary boundaries" that regulate the stability and resilience of the Earth (Stockholm Resilience Centre, 2017). As crucial planetary boundaries that have been crossed as a result of human activity, the Stockholm Resilience Centre (2017) pointed to climate change, loss of biosphere integrity, land-system change and altered biogeochemical cycles.

There is a wide range of environmental factors that affect investors' decisions and various classifications of them. However, in general, environmental issues might be divided in three broad groups (CFA Society of the UK, 2021: 97–114):

— climate change – this refers to changes in the composition of the global atmosphere caused directly or indirectly by human activity. Climate change has both local impacts (extreme weather events) and global impacts (the increase in global temperature and sea levels). Greenhouse gas emissions, especially carbon dioxide (CO_2), are indicated as the major drivers of climate change. To mitigate climate changes or adapt to them, sustainable solutions across various areas, including energy, buildings, transport, land use, agriculture, carbon pricing and manufacturing are necessary.

— pressures on natural resources – this encompasses progressive decline in biodiversity and less secure access to natural resources, such as fresh water, nutrient food, marine resources and other organic substances and raw materials. The reasons for the pressures are global population growth, economic development and increased consumption. Conserving nature and the improvement in sustainable resource use is crucial to safeguard biodiversity and, consequently, human health.

— pollution, waste and circular economy – this pertains to the air, water and soil pollution caused by human activities related to burning fossil fuels, excessive waste production or inappropriate waste management, and contamination of oceans, rivers or groundwater. Not only does it threaten biodiversity and destroy ecosystems but also impacts human health and wellness. A circular economy is an alternative approach to the use-make-dispose economy. This aims to minimise waste and preserve natural resources by designing a system which keeps products and materials in use for as long as possible while regenerating natural resources (Ellen MacArthur Foundation, 2019).

The economic costs of climate change have been estimated by numerous organisations pointing to significant losses. Alarmingly, these estimates are getting higher and higher every year. One report by the Economist Intelligence Unit (2015) estimated the net present value costs of climate change to achieve the level of USD 4.2 trillion by 2100. Then, according to IPCC (2018), the damage caused by climate change would be between USD 54tn and USD 69tn, depending on the scenario (1.5°C vs. 2°C increase in the global temperature). Munich Re (2020) estimated only the inflation-adjusted losses from extreme weather conditions such as flooding, droughts and storms in 2020 to be over USD 200bn. Undoubtedly, these figures justify the inclusion of environmental factors in the analysis of any investment process. Not surprisingly, at present the environmental dimension, and in particular climate and carbon footprint issues, remain a central focus for investment managers (Index Industry Association, 2022).

2. 3. 2 Social Factors

Social factors encompass issues which are significant to the lives of humans. This dimension refers to numerous aspects of human life including management of human capital, product liability and social opportunities. It focuses on the relationships between companies, employees, financial stakeholders, communities and policymakers. Examples of social factors are human rights, labour standards, health and safety, customer responsibility and animal welfare.

To assess investment opportunities and risks comprehensively, it is necessary to consider long-term social changes, also known as megatrends. They include the following issues (CFA Society of the UK, 2021: 192–206):

- Globalization – the process of increased global interactions resulting in the growth of international trade, and the exchange of culture and ideas. It might be perceived both as a risk for local communities and an opportunity leading to increased market efficiency. The globalisation process implicates other phenomena such as offshoring and dependency. Offshoring is based on the transfer of activities or whole industries to countries with lower wages and costs. Dependency relates to lowered country's independence in certain aspects and depending on foreign suppliers specialised in the particular field.
- Automation and artificial intelligence (AI) – the process of performing activities with minimal human assistance. On the one hand, it leads to faster production, lowered labour costs and the replacement of monotonous physical work by machines. On the contrary, it constitutes a risk associated with unemployment.
- Inequality and wealth creation – refer to growing discrepancies between rich and poor countries. According to the OECD Centre for Opportunity and Equality (Keeley, 2015), the average income of the richest 10% of the population is about nine times that of the poorest 10% across the OECD. This results in reduced health and education opportunities, not to mention the purchasing power among the lower and middle classes, and consequently, induces the total World's economic growth limitation.
- Digital disruption, social media and access to electronic devices – it refers to changes associated with the occurrence of new technological devices which affect the value proposition of existing goods and services. Depending on the industry, it might represent both an opportunity and risk.
- Changes to work, leisure time and education – the number of average working hours has decreased, while the level of education has increased. There are significant structural changes in the labour market resulting from automation and other factors. In corporations relying heavily on employees as a key asset, the assessment of human capital management strategies is vital.

— Changes to individual rights and responsibilities and family structures – more and more women are participating in the labour market; however, they are still less likely than men to have good quality jobs and relatively often face wage gaps.

— Changing demographics, including health and longevity – because of improvements in health care systems and changes in the lifestyle of the population in developed countries, people on average live longer but fewer children are born. An ageing population represents a major social risk affecting the workforce, national tax revenues, pension systems and some industries, including healthcare.

— Urbanisation – more and more people are living in urban areas. This constitutes some business opportunities related to the growing demand for infrastructure, but also risk deriving from pollution, waste management problems and unemployment among the less-skilled working class.

— Religion – it affects consumers' and investors' preferences.

Furthermore, there are some environmental megatrends which affect social issues. They include climate change and the related transition risk, water scarcity, mass migration, pollution and degradation of natural resources. These challenges can lead to an increase in the number of environmental migrants due to the inability to live in certain areas. For instance, Meze-Hausken (2000) investigated the vulnerability of people living in dryland areas focusing on migration in Northern Ethiopia. Then, Yan and Qian (2004) explored the issue of resettlements from the region of the upper Yangtze in China resulting from the degraded environment. Finally, conducted a review of studies and showed that out of 31 empirical articles, 23 delivered some evidence of the influence of environmental factors on migration across international borders.

Having discussed the principal global social trends, it is time to consider the social factors that determine investors' decisions. Basically, social factors can be divided into two groups: internal factors that affect employees and external factors that affect local communities, customers and other stakeholder groups.

To begin with internal factors, human capital development is the first to be identified. This refers to a human capital management system that ensures high labour productivity, continuously improving employees' competencies while keeping them engaged and motivated. Moreover, this factor encompasses social inclusion and active citizenship to improve employability. According to Kotsantonis and Serafeim (2020), because of the rapid technological improvements and the uncertainty of future work, the importance of human capital is likely to increase.

Another widely recognised social factor includes working conditions, health and safety. The prevention of workforce accidents and fatalities and minimising the risk of occupational diseases is a central focus of this factor. Then, human rights include, for instance, the right to life and liberty, freedom from slavery and torture, freedom of opinion and expression, and the right to work and education. Human rights are globally regulated by the Universal Declaration of Human Rights (UN, 1948) and, regarding businesses, by the United Nations Guiding Principles on Business and Human Rights (Business and Human Rights Resource Centre, 2019) and OECD Guidelines for Multinational Enterprises (OECD, 2023). Last, but not least, labour rights aim to ensure decent and productive work both for men and women, as well as the right to organise and collective bargaining, providing a living wage and equal enumeration. Labour rights are to combat forced labour, worst forms of child labour and occupational discrimination (CFA Society of the UK, 2021: 198-204).

Then, external factors pertain to stakeholder opposition, controversial sourcing, product liability, consumer protection, social opportunities or animal welfare. Among the social factors, good relations with the community play a vital role, as a company's unethical behaviour often leads to conflicts with local communities and reputational damage, posing a significant risk to the company's growth. Similarly, keeping consumers satisfied, and providing safe products free of manufacturing defects is essential for building customer trust and a solid brand reputation. Importantly, McKinsey research conducted by Henisz, Koller and Nuttall (2019) showed that 70% of consumers surveyed on purchases in multiple industries would pay an additional 5% for a green product if it met the same performance standards as a nongreen alternative. This proves that companies following ESG standards benefit from it. The support of social initiatives such as education, healthcare or access to water has no longer been attributed only to charities, but also to companies. Consequently, it is becoming more and more common for businesses to make a real positive contribution to society (CFA Society of the UK, 2021: 204–206).

Comprehensive investment analysis requires the identification of key material country, sector and company factors and social trends. Depending on the legal and cultural environment in which the company operates, or the level of economic development of the country, different factors will be more or less important. In the textile sector, the protection of children's rights is of great importance, while in the high-tech sector, the development of artificial intelligence is crucial. At the company level, on the other hand, the investor should pay attention to the organisational culture, the protection of workers' rights or the compatibility of the company's strategy with consumer expectations (CFA Society of the UK, 2021: 206–209). Overall, the multitude of factors and ways of measuring the social pillar of ESG prove how challenging the process is. Not only does it consider the specificity of the sector, but also other factors influencing the investment decision.

Finally, it should be highlighted that due to the wide array of factors encompassing the social dimension of ESG, there are significant misalignments in social scores between various ESG rating providers (Chatterji et al., 2016). Depending on the data used and the scope of analysis, ESG ratings might differ. This is why scholars commonly report divergences in ESG scores and, in consequence, low comparability (Delmas, Etzion and Nairn-Birch, 2013). This is why the importance of the social dimension of ESG is commonly undervalued by investment managers. As pointed out in one report by Index Industry Association (2022), challenges and difficulties associated with measuring intangible aspects of ESG, including social issues, lead to diminished attention on these factors analysis as compared to environmental factors which seem to be more tangible.

2. 3. 3 Governance Factors

Corporate governance is a broad term denoting a system of management and oversight of an organisation that, through the application of appropriate policies and rules, ensures the cohesiveness of an organization. In essence, this system encompasses all the processes in the organisation and the people involved in their execution (principally the individuals on the board) in order to achieve strong company performance while mitigating risk and optimising internal efficiency. According to Oman (2001), corporate governance refers to both private and public institutions that include laws, regulations, and business practices that govern the relationship between corporate managers and stakeholders. Likewise, La Porta et al. (2002) highlight that corporate governance encompasses a set of mechanisms that enable outside investors (shareholders) to protect themselves from internal investors (managers).

Successful corporate governance enhances a company culture underpinned by integrity and honesty, contributing to the sustainability of the entire organisation to the benefit of shareholders and the company's other stakeholders (CFA Society of the UK, 2021: 228). Importantly, Yu, Luu, and Chen (2020) evidenced that good corporate governance practices are essential in preventing greenwashing behaviours among companies. They argued that the incorporation of independent directors, institutional investors, influential public interests via a less corrupt country system, and cross-listings of companies affect positively a company's disclosure ethics. Finally, the broad scope of corporate governance issues (see Figure 10) disclosed by companies contributes to the improvement of environmental and social factors reporting as well (World Economic Forum, 2022).

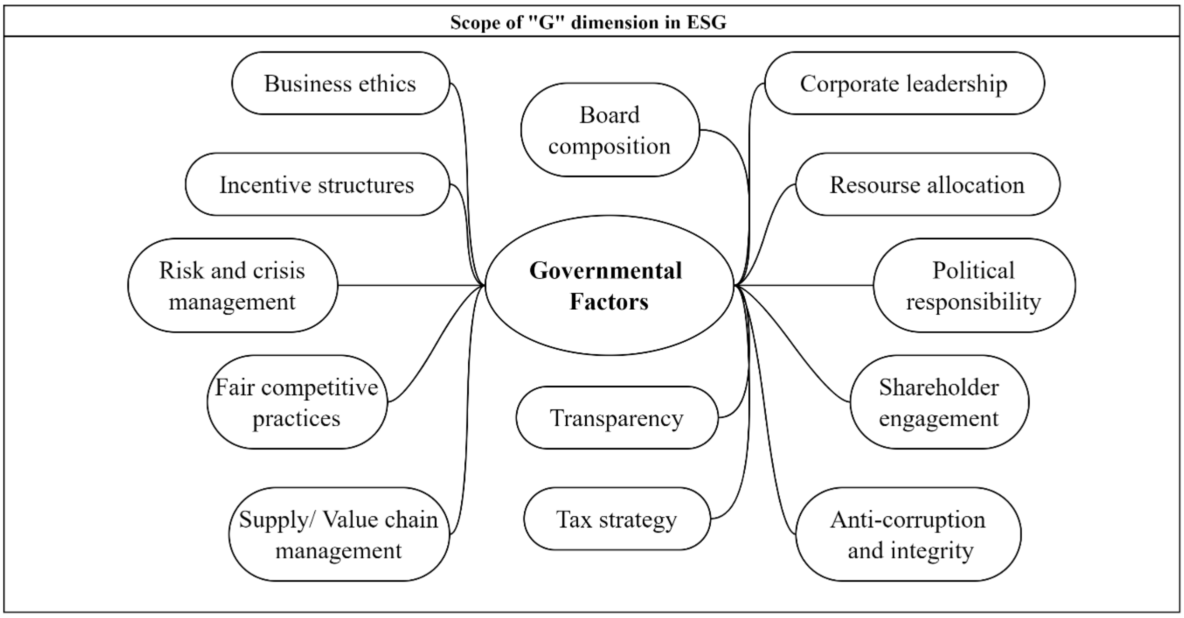

Figure 10. Scope of governmental factors
(Source: Own elaboration based on World Economic Forum, 2022)

The regulations, principles, and norms that comprise corporate governance come down to two key concerns: accountability and alignment. Accountability refers to providing people with authority and responsibility for the tasks they are given. It is necessary to link their actions to results and ensure that they bear the consequences of their decisions. The problem of accountability is particularly evident among top management. For instance, Becher and Campbell (2004) examined the corporate governance of bank mergers and acquisitions and showed that CEOs negotiate for their self-interest while the company's external directors struggle financially. Therefore, corporate governance attaches great importance to the independence and appropriate structural diversity of directors to avoid groupthink or over-concentration of power. According to Kamalnath (2018), gender diversity might relieve the problem of groupthink only if women stay independent and bear the “outsider” status. If women become part of the “in-group”, they may stop thinking independently and questioning decisions, which can weaken their impact on how the board makes choices.

Then, alignment pertains to the agency problem. It refers to the discrepancies between the interests of professional managers and the owners of the business. The problem of agency costs has been known for a long time. Indeed, Smith (1776) already suggested the existence of a conflict of interest between business owners and managers. Farinha (2003) pointed out that managers and shareholders have fundamentally different approaches and objectives for the company's operations. As the manager's perspective is associated with limited cash flows, managers focus on a short-term investment perspective, while shareholders are primarily concerned with a quick return on investment. Along the same lines, the two groups differ in their risk preferences. Shareholders focus on market risk, while managers are mainly concerned with company risk, as their employment depends on the company's risk. In addition to the causes of the agency problem mentioned earlier, Chowdhury (2004) pointed to the short duration of agents' involvement in the organisation, unsatisfactory incentive plans for agents, and the presence of information asymmetry within the company.

One more conflict of interest may arise between the majority owners and minority owners. As noted by Fama and Jensen (1983), the majority owners have more voting power and can force all decisions in their favour while harming the interests of minority shareholders. This problem mainly affects companies in which ownership concentrates in the hands of a few individuals or family owners, and minority shareholders find it difficult to protect their interests or wealth (Demsetz and Lehn, 1985). Then, Ilhan-Nas et al. (2018) argued that family ownership owners apply internal corporate governance and informal relationships to achieve their individual goals.

All in all, corporate governance factors are pivotal in the functioning of any modern ESG-oriented company. Not only does it contribute to the increased care of all the stakeholders' interests, but it also induces better implementation of environmental and social factors disclosure. However, similarly to the social dimension, governance factors feature a high level of complexity and often intangibility, which commonly blocks investment managers' interest in the thorough analysis of the "G" dimension (Index Industry Association, 2022). A better understanding and systematisation of ESG issues is intended to be achieved through widely undertaken ESG initiatives, which are the subject of the subsequent subsection.

2. 4 ESG Initiatives

International organisations, their initiatives, interpretations, and guidelines have played a huge role in developing and spreading the concept of ESG investing. Not only have they made investors aware of the importance of incorporating environmental, social, and governance factors into the investment process, but they have also provided the necessary information and guidance on the right way to understand, measure, and analyse ESG issues. The following section will introduce the key bodies and their initiatives addressing ESG investing matters.

2. 4. 1 United Nations Initiatives

One key organisation that has contributed significantly to the dissemination of responsible investment is the United Nations. These days, the world's largest international initiative focusing on sustainable business is the United Nations Global Compact (UNGC). It was established in 2000 by the UN Secretary-General as a result of cooperation between the UN and leading corporations. The initiative is based on voluntary CEO commitments to support sustainable development. In 2022, it already had more than 10,000 signatories from all over the world. The United Nations Global Compact strives to mobilise a global movement of sustainable companies and stakeholders to create a safe, fair, and human-friendly world. All the signatories agree to adhere to the UNGC's ten principles concerning human rights, labour, environment, and anti-corruption. They are derived from the broader standards such as the Universal Declaration of Human Rights and the International Labour Organization’s Declaration on Fundamental Principles and Rights at Work. The fundamental aim of these principles is to provide investors with guidance on how to assess and engage with businesses, as well as meet broader social goals such as the UN Sustainable Development Goals.

Another crucial initiative is the United Nations Environment Programme Finance Initiative (UNEP FI). UNEP FI was established in 1992 as a partnership between the UNEP and the global financial sector to mobilise private sector finance for sustainable development. In 2022, the initiative has already encompassed the UNEP’s cooperation with over 450 banks, investors, and insurers. The UNEP FI has developed or co-created the frameworks for the incorporation of sustainable development principles into the financial industry, including the Principles for Responsible Investment (PRI), the Principles for Sustainable Insurance (PSI), and the Principles for Responsible Banking (PRB).

The PRI is the leading international organisation of investors working for responsible investment. The initiative was launched in 2006 through a collaboration between UNEP FI and the UN Global Compact. While initially, the PRI comprised only a few prominent investors, by 2022, it united more than 4,000 signatories collaborating to develop clear and understandable practices for the ESG incorporation. The PRI's activity centres on four aspects:

— providing a wide range of ESG tools and good practices for managers and equity owners, as well as consultants and providers of ESG information.

— supporting investors' engagement in companies and sectors consistent with responsible investment principles.

— reviewing, analysing, and consulting ESG-related policies and supporting communication between investors and regulators.

— the PRI Academy develops, aggregates, and disseminates academic research on responsible investment.

The PRI provides guidance on how to incorporate ESG issues into the investment process. There are six voluntary PRI principles:

1. We will incorporate ESG issues into investment analysis and decision-making processes.
2. We will be active owners and incorporate ESG issues into our ownership policies and practices.
3. We will seek appropriate disclosure on ESG issues by the entities in which we invest.
4. We will promote acceptance and implementation of the Principles within the investment industry.
5. We will work together to enhance our effectiveness in implementing the Principles.
6. We will each report on our activities and progress towards implementing the Principles.

Then, from the environmental viewpoint, one crucial initiative is the United Nations Framework Convention on Climate Change (UNFCCC). It was established at the Rio de Janeiro Earth Summit in 1992 to support the reduction of greenhouse gas emissions and mitigate man-made climate change. The UNFCCC is a host of annual Conferences of the Parties (COP) which aim to advance its member states' voluntary commitments to climate change limitation. Among all the COP meetings, two were of especially great importance: the COP3 meeting in Kyoto (1997) and the COP21 meeting in Paris (2015). The result of the Kyoto meeting was the Kyoto Protocol, which committed industrialised countries to reducing GHG emissions. The COP21 meeting, in turn, has led to the signing of the Paris Agreement, which committed developed and emerging countries to strengthen the efforts of climate change mitigation by keeping the global temperature rise in the 21st century well below 2°C above the pre-industrial level.

Finally, the UN Sustainable Development Goals (SDGs) were approved by all the UN members in 2015. The SDGs, which replaced the UN Millennial Goals, comprise 17 global challenges of key importance for humans and the planet. They include the urgent problems associated with poverty, inequality, climate change, environmental degradation, peace, and justice. Every year, the UN Secretary-General presents the Sustainable Development Goals Progress Report, which provides key indicators and data related to SDGs implementation at the national and regional levels. Table 5 presents the consecutive Sustainable Development Goals from the 2030 Agenda (United Nations, 2015). Importantly, the table delivers just example indicators since the total number of individual indicators of SDGs proposed by the UN is 230 (United Nations Secretary-General, 2016).

Table 5. Sustainable Development Goals with selected indicators

Nr	Goal	Indicators
1	End poverty in all its forms everywhere	• Proportion of population below the international poverty line, by sex, age, employment status, and geographical location (urban/rural)
2	End hunger, achieve food security and improve nutrition, and promote sustainable agriculture	• Proportion of agricultural area under productive and sustainable agriculture • Volume of production per labour unit by • classes of farming/pastoral/forestry enterprise size
3	Ensure healthy lives and promote well-being for all at all ages	• Maternal mortality ratio • Death rate due to road traffic injuries • Total net official development assistance to medical research and basic health sectors
4	Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all	• Participation rate of youth and adults in formal and non-formal education and training in the previous 12 months, by sex • Percentage of population in a given age group achieving at least a fixed level of proficiency in functional (a) literacy and (b) numeracy skills, by sex
5	Achieve gender equality and empower all women and girls	• Proportion of time spent on unpaid domestic and care work, by sex, age, and location • Proportion of women in managerial positions
6	Ensure the availability and sustainable management of water and sanitation for all	• Proportion of population using safely managed drinking water services • Proportion of wastewater safely treated
7	Ensure access to affordable, reliable, sustainable, and modern energy for all	• Proportion of population with access to electricity • Renewable energy share in the total final energy consumption • Energy intensity measured in terms of primary energy and GDP
8	Promote sustained, inclusive and sustainable economic growth, full and productive employment, and decent work for all	• Annual growth rate of real GDP per capita • Proportion of informal employment in non-agriculture employment, by sex
9	Build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation	• Proportion of the rural population who live within 2 km of an all-season road • Research and development expenditure as a proportion of GDP
10	Reduce inequality within and among countries	• Proportion of people living below 50 per cent of median income, by age, sex and persons with disabilities • Number of countries that have implemented well-managed migration policies
11	Make cities and human settlements inclusive, safe, resilient and sustainable	• Proportion of urban population living in slums, informal settlements or inadequate housing • Ratio of land consumption rate to population growth rate
12	Ensure sustainable consumption and production patterns	• Material footprint, material footprint per capita, and material footprint per GDP • Global food loss index
13	Take urgent action to combat climate change and its impacts	• Number of countries with national and local disaster risk reduction strategies

		• Number of countries that have integrated mitigation, adaptation, impact reduction and early warning into primary, secondary and tertiary curricula
14	Conserve and sustainably use the oceans, seas, and marine resources for sustainable development	• Index of coastal eutrophication and floating plastic debris density • Coverage of protected areas in relation to marine areas
15	Protect, restore, and promote sustainable use of terrestrial ecosystems, sustainably manage forests, combat desertification, and halt and reverse land degradation and halt biodiversity loss	• Proportion of important sites for terrestrial and freshwater biodiversity that are covered by protected areas, by ecosystem type • Proportion of land that is degraded over total land area • Proportion of traded wildlife that was poached or illicitly trafficked
16	Promote peaceful and inclusive societies for sustainable development, provide access to justice for all, and build effective, accountable, and inclusive institutions at all levels	• Number of victims of intentional homicide per 100,000 population, by sex and age • Proportion of the population that feels safe walking alone around the area they live in • Proportion of the population satisfied with their last experience of public services
17	Strengthen the means of implementation and revitalize the Global Partnership for Sustainable Development	• Total government revenue as a proportion of GDP, by source • Developing countries and least developed countries' share of global exports • Number of United States dollars committed to public-private and civil society partnerships

Source: Own elaboration based on United Nations Secretary-General (2016)

2. 4. 2 Reporting Initiatives

Another extremely important issue determining the development of ESG investments is the reporting of environmental, social, and governance issues. Unfortunately, there is no single universal standard for ESG reporting yet, so various, mostly voluntary ESG frameworks are used at the company, regional, and national levels. This section of the paper will present the most relevant contemporary ESG reporting initiatives applied both regionally and internationally.

Beginning with one of the most well-known ESG reporting guidelines, one should refer to the Global Reporting Initiative (GRI), founded in 1997. What makes the initiative special is that *GRI standards* comprise comprehensive guidance on disclosure across E, S, and G factors for all shareholders, not only investors, as the majority of global ESG frameworks. The Global Reporting Initiative is applied by several thousand organisations all over the world, including the United Nations Global Compact.

While GRI standards are distinguished by their broad multi-level ESG disclosure, the Sustainability Accounting Standards Board (SASB) was established in 2011 to provide industry guidance from a financial perspective. The SASB aims to identify the subset of environmental, social, and governance issues crucial to financial performance in each of the 77 industries. It constitutes principles and standards to guide companies on how to identify, manage, and communicate financially relevant sustainability information to investors. In 2020, GRI and SASB merged forces and next published a guide on how organisations can apply both standards simultaneously (GRI and SASB, 2021).

Then, the Value Reporting Foundation (VRF) was formed in 2010 upon the merger of the International Integrated Reporting Council (IIRC) and the Sustainability Accounting Standards Board (SASB) to provide investors and organisations with a comprehensive guide on an enterprise's value drivers. The Foundation aimed to help corporations develop, manage, and communicate a strategy that creates long-term value and contributes to improved performance.

The Climate Disclosure Standards Board (CDSB) was established in 2007 to offer companies a framework for reporting environmental and social information with the same rigour as financial information. Furthermore, the CDSB guidelines formed the basis for the recommendations of the Task Force for Climate-Related Financial Disclosures (TCFD), whose mission is to improve and increase the reporting of climate-related financial information.

In 2022, there was a consolidation of the Value Reporting Foundation (VRF) and the Climate Disclosure Standards Board (CDSB) into the International Financial Reporting Standards Foundation (IFRS Foundation). The foundation work to develop a comprehensive global baseline of sustainability disclosures for the capital markets. Now, the IFRS Foundation's resources include the Integrated Thinking Principles, the Integrated Reporting Framework, and the SASB Standards.

On the regional level, one most comprehensive ESG reporting standards is set by the EU's Sustainable Finance Disclosure Regulation (SFDR). It was introduced to enhance transparency and prevent greenwashing in the sustainable finance market. Ultimately, the initiative aims to impose comprehensive sustainability disclosure requirements covering a wide range of ESG indicators at both the entity and product levels. The initiative supports the European Green Deal - a set of policy measures to tackle the climate crisis by transforming the EU into a modern, resource-efficient, and competitive economy with no net greenhouse gas emissions by 2050.

As demonstrated above, there is still a lack of standardisation regarding ESG reporting. This, in turn, has negative repercussions on data integrity and comparability. However, the increasing cooperation of peer initiatives, joining forces, and the development of common standpoints can be pointed out as a positive trend in recent years. Hopefully, in the future, global cooperation will allow for the establishment of universal standards for ESG measurement and reporting worldwide.

2. 5 Investment Performance and Risk of ESG Investments

A fundamental question of any investment strategy is how it affects the value of an investor's portfolio. With the growing popularity of sustainable investing, investors and academics have started to investigate the risk-return characteristics of ESG portfolios. The following part of the thesis will present some academic research on the investment results of ESG investments.

2. 5. 1 Theoretical Considerations

One essential paper on the impact of SRI strategy on investment performance was *The Wages of Social Responsibility* (Statman and Glushkov, 2009). The authors proposed three alternative hypotheses about the return formation on investments in socially responsible companies. First, – *doing good but not well* states that the expected rate of return on responsible investment is lower than the expected rate of return on conventional investment. This is because a company's application of responsible investment strategies incurs high costs that are not compensated by an increase in investment performance. A decrease in a company's profits, in turn, results in a decrease in the return on investment of the instruments it issues (Statman and Glushkov, 2009: 36). This thesis is supported by numerous studies, including a meta-analysis of 167 studies which indicated that excessive or poorly managed socially responsible initiatives can negatively impact profitability and investment returns. Barnea and Rubin (2010) showed that managers engage in socially responsible activities whose costs outweigh the benefits because they are motivated by personal gain.

According to the concept of asset selection costs, socially responsible investments show lower efficiency than conventional investments due to the generation of additional costs. Analysing an investment from the view of environmental, social, and governance issues involves accessing additional information, which requires managers more time and work. This, in turn, contributes to reduced returns for the investor. The study on the funds' expense ratio (Morningstar, 2021) evidenced that sustainable funds have a higher asset-weighted average expense ratio (0.61% at the end of 2020) than their traditional counterparts (0.41%).

Another hypothesis – *doing good while doing well* – is based on the assumption that expected returns on responsible investment are higher than expected returns on conventional investment due to managers and investors underestimating the benefits of a company's socially responsible activities or overestimating the costs of such activities (Statman and Glushkov, 2009: 36). This thesis is supported by research by Lev, Sarath and Sougiannis (2005), who noted that investors focus on a company's reported profitability and underestimate the benefits of R&D activities contributing to increased profits over the long term. Similarly, Edmans (2010) remarks that managers may overlook the significance of intellectual capital such as employee satisfaction, which, despite reducing a company's short-term profits, provides material benefits in the long run.

The third hypothesis - *no effect* - states that the expected returns on investments in shares of socially responsible companies and those of conventional companies do not show significant differences. This hypothesis may be true in two cases. Firstly, when a company's CSR activities do not generate any additional costs, and secondly, when the costs of CSR are offset by the benefits of CSR (Statman and Glushkov, 2009: 39).

From the perspective of the Modern Portfolio Theory (1952), a responsible investing strategy must decrease investment returns or increase investment risk as the set of investment opportunities decreases. This means that a portfolio comprising sustainable equities will have lower risk-adjusted performance than a traditional portfolio including all assets in the market. Screening out companies into a sustainable portfolio reduces diversification opportunities (as compared to investing in a broad market index). Thus, any investment portfolio different from a broad portfolio of assets (including the sustainable one) provides a lower return for a given level of risk or a higher risk for a given level of return.

However, there is no shortage of opinions in both financial practice and academia that contest the validity of Markowitz's portfolio theory. Specifically, the behavioural finance theory based on two theories of choice under uncertainty contradicts the Modern Portfolio Theory assumptions. Firstly, the SP/A theory (security-potential/aspiration) developed by Lopes (1987) applies the aspiration level as a second criterion in the choice process. Subsequently, the prospect theory (Kahneman and Tversky, 1979) indicates that investors value gains and losses differently, attaching more weight to perceived gains than perceived losses. While mean-variance investors choose portfolios by considering mean and variance, behavioural portfolio theory investors choose portfolios by considering expected wealth, desire for security and potential, aspiration levels, and probabilities of achieving aspiration levels (Shefrin and Statman, 2000: 128). Opponents of Modern Portfolio Theory argue that it is based on impossible assumptions, as all investors are rational and have the same expectations of return and risk on investments. In reality, investors often prove to be irrational, and sustainable investing is one of those investment approaches in which investors consider non-financial factors that may not carry an increased rate of return or reduced risk. Non-financial motives, such as the desire to support pro-environmental and pro-social enterprises, would have to be considered irrational, whereas this can simply be a broader and more long-term dimension of rationality that considers the welfare of future generations.

2. 5. 2 Sin Stocks Performance

Sin stocks are the type of assets that the underlying businesses are unacceptable to the majority or part of society due to profiting from human weaknesses and frailties. Sin stock sectors include industries related to alcohol, tobacco, gambling, sex, arms manufacturing, and the military, but this term may be understood slightly differently by particular investors. There is strong evidence in the scientific literature on the sin stocks' outperformance over green stocks and the broad market. It was Merton (1987) who first indicated that when a large group of investors neglect certain stocks, for instance, low-ESG stocks, they become undervalued. Despite initial low returns, in the long run, those stocks provide high returns relative to high-ESG stocks. Low stock valuation implies a high dividend/price ratio, and consequently, higher returns, so the long-term undervaluation brings abnormal profits. The following part of the dissertation deals with the key research on the topic of sin stocks' outperformance.

The price of sin: The effects of social norms on markets by Hong and Kacperczyk (2009) is arguably the most well-known article on the return effects of negative screening. The Authors found that sin stocks (defined in the paper as tobacco, alcohol, and gambling businesses) are commonly ignored by institutional investors and financial analysts relative to the control group of stocks. The study covered the years 1926–2006 and showed that sin stocks exhibited an annual outperformance of approximately 3–4% relative to comparable non-sin stocks. Despite some criticism of the research methodology (Hoepner and Zeume, 2013; Adamsson and Hoepner, 2015), the paper remains the most prominent and most cited.

Then, Richey (2017) investigated whether a portfolio of vice stocks (precisely stocks of companies manufacturing or selling products such as alcohol, tobacco, gaming services, and national defence) performs better than the market portfolio (S&P 500 index). Employing several factor models, Richey examined 61 corporations from vice-related industries in the period October 1996 to October 2016 and found that the S&P 500 returned annually 7.8% while the Vice Fund returned 11.5%. What is more, the Vice Fund exhibited lower market risk (its portfolio beta was between 0.59 and 0.79 while the S&P 500 Index had a beta of 1). The annual returns of sin portfolios above the risk-adjusted benchmark (alphas) on the CAPM, three-factor, and four-factor models were 2.9%, 2.8%, and 2.5%, respectively, and were statistically significant at the 0.01 level. However, in the five-factor model (which adds the investment and profitability factors), the annual alpha of vice portfolios fell to 0.1%. The Author suggests that higher risk-adjusted returns of vice stocks result from a more profitable and less wasteful operation than the average company.

A new piece of evidence on the outperformance of sin stocks was provided by Dimson, Marsh, and Staunton (2020). They investigated industry indices in the period 1900-2019 and found that the highest returning industries were tobacco and alcohol. While the broad U.S. stock market brought an average annual return of 9.6%, the tobacco sector returned 14.2%. The U.K. stock market in the same period returned on average 9.3%, while the alcohol industry gave an annual return of 11.5%. The results are consistent with the economic theory as investors demand excessive returns for risky assets.

Finally, Zerbib (2022) proved that exclusionary screening and ESG integration impact asset returns. He constructed the Sustainable Capital Asset Pricing Model (S-CAPM) with a premium for neglected stocks and a taste premium. Then, he applied the model to green investing and sin stock exclusion. The sample encompassed 348 green funds investing in U.S. equities between 2000 and 2018. Zerbib showed that the S-CAPM model outperformed the Carhart four-factor model, and the model yielded a 1.5% taste premium and 2.5% exclusion premium per annum. What is more, with an increase in investors' pessimism about the particular asset, the premium increases.

2. 5. 3 ESG and Risk-adjusted returns

A great number of present studies focus on the relationship between an asset's ESG score and the expected risk-adjusted returns on investment. Therefore, the following section will explore fundamental research dealing with that issue.

Ciciretti, Dalò, and Dam (2017) noted that the demand for SRI funds might be explained by the favourable risk characteristics of responsible assets and investors' taste for such assets. To investigate the taste effect's contribution to risk-adjusted returns, they build a model considering the market beta, size, value, momentum, and SRI score factors. The SRI score was calculated based on six dimensions, including business behaviour, corporate governance, community involvement, environment, human resources, and human rights. The study was conducted in the period July 2005 through June 2014 and covered 1000 corporations from the U.S., Europe, and the Asia-Pacific region. They found that with an increase in the portfolio SRI score, the average monthly excess return declined. There was a significant and negative relationship between social responsibility scores and risk-adjusted returns. The “price of taste” amounted to 4.8% annually, and the excess return deriving from the strategy that buys the portfolio with the worst SRI score and sells the portfolio with the best SRI score was 7.2 percentage points annually. The Authors concluded that investors with SRI preferences pay a price in terms of lower returns.

Another study by Ciciretti et al. (2023) confirmed that firms with lower ESG scores exhibit higher expected returns. This is due to an increased risk associated with investing in corporations with low ESG scores. Investors who accept the excessive risk demand increased returns, and businesses with low ESG scores pay a price in the form of a higher cost of capital. The Authors build an equilibrium model and provide a piece of evidence that the cross-sectional variation of expected returns associated with the ESG premium is principally driven by investors’ preferences rather than the systematic risk components captured by ESG scores.

The same conclusions were reached by Pastor et al. (2021), who supported the thesis that with the increased demand for ESG funds, their expected long-term returns fall. Green assets exhibit negative alpha while brown assets have a positive alpha. Generally, the greener the asset, the lower the CAPM alpha in equilibrium. However, green stocks might deliver short-term higher returns than brown stocks, resulting from the increased demand for green investments. What is more, green assets are more exposed to the ESG risk factor. With the strengthening concerns about ESG issues, the demand for green products and services rises, and, consequently, investors might derive more utility from holding green stocks. The premium related to the ESG risk factor might be large enough to overcome the negative alpha of green investments.

Winegarden (2019) conducted a study on the performance of the relatively few sustainable mutual funds with a long listing history. He examined 30 ESG funds that either existed for more than 10 years or had outperformed the S&P 500 Index over a short period. Having constructed a USD10,000 ESG portfolio (equally divided across the funds and considering the management fees), the Author found that after 10 years, the value of the ESG portfolio would be 43.9% smaller than the investment in the S&P 500 Index fund. Over the 10-year horizon, only two sustainable funds exceeded the returns of a benchmark, and over the 5-year horizon, only one fund outperformed the S&P 500 benchmark. Winegarden concluded that, as far as the performance is concerned, ESG funds have failed to match the performance of a simple, broad-based index fund yet.

Then, Plagge and Grim (2020) achieved mixed results on the ESG equity funds' performance. They investigated the American-market-oriented index and active funds, including ETFs which incorporate ESG strategy. The study covered the years 2004-2018, with the dataset including 98 funds at the beginning and 267 funds at the end of the period. The Authors showed that after controlling for style factor exposures, the majority of ESG funds did not produce any statistically significant positive or negative gross alpha. Furthermore, they noted that the differences in risk-return characteristics between the individual ESG funds were principally driven by fund-specific criteria (such as large-cap or mid-cap focus or industry concentration) rather than by a homogenous ESG factor. Lastly, they evidenced that over time the positive alphas for some active funds have declined, indicating a growing market efficiency. Overall, Plagge and Grim suggested a careful assessment of individual ESG funds rather than following generalizations on the risk and return characteristics of these funds.

Another crucial issue related to ESG investing is the impact of environmental, social and governance screening on investment performance. According to the portfolio theory, any screening deteriorates the risk-return characteristics relating to the market portfolio, nevertheless, as the review demonstrates, historically ESG screening has delivered mixed results.

Derwall et al. (2005) examined the returns related to the strategy based on Innovest Strategic Value Advisors' corporate eco-efficiency in the years 1995–2003. Their study showed that more eco-efficient firms exhibit higher stock returns than their less eco-efficient counterparts. However, as the investigated period was relatively short, the results could not be perceived as robust. Several years later, Derwall et al. (2011) correlated the Innovest eco-efficiency data with operating performance measures and equity valuation. They proved that during the sample period eco-efficient firms become relatively more expensive, which implies that the return outperformance of the 2005 study was associated with the changes in valuation. It means that either eco-efficient firms were initially undervalued or that over time they became overvalued, and the abnormal returns of investment in eco-efficient companies were just temporary market inefficiency.

The results of studies on the performance of an environmentally conscious investment strategy based on KLD scores are quite optimistic for ESG investors. Kempf and Osthoff (2007) found that over the period from 1991 until the end of 2003, buying stocks with high socially responsible ratings and selling stocks with low socially responsible ratings brought abnormal returns of up to 8.7% per year. In addition, they remarked that the maximum abnormal returns were linked with the best-in-class screening approach and the use of a combination of several socially responsible screens simultaneously. Then, Halbritter and Dorfleitner (2015) indicated that a long-short portfolio yielded a four-factor alpha of 6.6% per year during 1990-2001 (at the 0.01 significance level), while over the years 2002-2012, the strategy delivered negative alphas (albeit these results were statistically insignificant).

Bolton and Kacperczyk (2020) examined the relationship between carbon emissions and stock returns in the U.S. market over the years 2005–2017. They categorized a company's carbon emissions into three scopes: direct emissions from production, indirect emissions from consumption of purchased electricity, heat, or steam, and indirect emissions from the production of purchased materials, product use, waste disposal, and outsources activities, etc. The study revealed that stocks with higher total carbon emissions earn higher returns after controlling for size, book-to-market, momentum, beta and liquidity factors, as well as net stock issuance. Furthermore, they found a significant carbon premium related to growths in carbon emissions. This is consistent with the economic theory, as firms successfully reducing CO_2 emissions can afford to offer lower returns, while more risky fossil fuel-based companies must offer higher returns to compensate for the additional risk.

As suggested by Pastor, Stambaugh and Taylor (2019), in the short term the performance of green assets may exceed the return of brown assets even though it is a brown asset that exhibits a higher expected return. The ambiguous relationship between carbon risk and short-term returns was investigated in the study *Decarbonisation Factors* (Cheema-Fox et al., 2021). They found that strategies associated with buying decarbonisation factors when institutional flows were positive and selling the factors when flows were negative brought significant outperformance. In the U.S. alphas provided by this strategy were between 1.5% and 4.4%, while in the European market between 2.5% and 8.5%.

When it comes to social screening, there is some evidence that the performance on investment is affected by at least one social screen, namely employee satisfaction. Edmans (2010) proved that companies with high employee satisfaction exhibit high future stock returns. The Author noted that the value-weighted portfolio of the 100 Best Companies to Work For in America provided an annual four-factor alpha of 3.5% from 1984–2009, and 2.1% above industry benchmarks. Then, Edmans (2012) repeated the research to cover the years 1984-2011 and achieved almost identical results. He argued that the above-average returns of companies with high employee satisfaction scores result from the market undervaluation of intangible information. Over time, intangible information becomes tangible through higher company returns and, consequently, the pricing of these companies (and thus investor returns) increases. The results are consistent with Kempf and Osthoff (2007) and Statman and Glushkov (2009) who showed that a strategy based on KLD scores on employee relations (and community) brings high returns.

Then, studies show that the relationship between a firm's governance score and investor returns has changed over time. Gompers, Ishii, and Metrick (2003) constructed a firm-level governance index (G-index) based on 24 provisions that weaken shareholder rights (a company with weak shareholder rights had a high G-index, and a firm with strong governance had a low G-index). For a sample of 1500 large U.S. firms during 1990–1999, they found that a portfolio holding the 10% lowest G-index firms and shorting the 10% highest G-index firms achieved an abnormal return of 8.5% per year. Subsequently, Bebchuk, Cohen and Wang (2013) conducted a similar study in the years 1990–2008 and evidenced that the abnormal returns are insignificant during the period 2000–2008. This suggests that the previously identified effect has disappeared after the original sample period.

However, neither the first-mentioned study nor the second found evidence that good governance firms are riskier. Therefore, the higher returns of high G-scored companies represent a market anomaly. Hvidkjær (2017) pointed out that, arguably, investors in the 1990s were not aware of the detrimental effects of the governance provisions, and in the 2000s, with the growing focus on governance issues, they had become conscious of the effects.

Auer (2016) examined the effect of exclusionary screening on portfolio Sharpe ratios. His study included the data of STOXX 600 (index of large-cap European stocks) over the years 2004-2012 and applied ESG ratings from Sustainalytics. The Author noted that the Sharpe ratio of the rated stocks increases when excluding stocks with poor governance rating (at the significance level of 0.05), while exclusionary screening based on environmental and social screening does not affect the Sharpe ratios.

One strategy to improve the profitability of sustainable investment is to purchase ESG companies with a high cost of capital. The Authors of the study *Outperformance through Investing in ESG in Need* (Hsu et al., 2018), evidenced that firms with a higher cost of equity outperform firms with a lower cost of capital. To evaluate the cost of capital, they applied such measures as book-to-market, gross profitability, net operating assets, accruals, volatility, asset growth, and market beta. The study showed that a high cost of capital ESG strategy outperformed the S&P 500 index by about 3 percentage points annually. In general, however, excessive returns resulting from ESG screening are not possible in the long term. The conclusion from the investigated studies is that, from time to time, ESG screening strategies might deliver alpha, but this is rather a short-term anomaly due to market inefficiencies rather than a stable regularity.

Modern research shows that the relationship between sustainability measures and expected returns is nuanced (Pedersen et al., 2020; Pastor et al., 2021). Depending on how investors behave, securities with a high ESG score will yield lower or higher returns than the benchmark. When investors avoid securities with poor ESG scores, they will generate low returns. However, in the future, they may generate high returns as compensation for investors' previous negligence. In turn, high capital flows into securities with high ESG ratings may produce abnormal returns at the time this net flow occurs, but once a new equilibrium of holdings is reached, these same securities may experience lower future returns (Pastor et al. 2021). This carries implications for the usefulness of research into the relationship between ESG performance and future rates of return, as depending on the period adopted, the author may achieve contrasting results. Bansal et al. (2022) argued that green stocks resemble luxury goods since there is a higher demand for them when the economy does well, and the financial concerns are weaker. Therefore, due to investors' preferences, green stocks might outperform brown stocks during the prosperity periods; however, during a bear market, when the financial risk is greater, they tend to underperform.

As the conclusions derived from various studies are dubious, it would be useful to focus on one meta-study. Whelan et al. (2021) reviewed over 1,000 studies published between 2015–2020. Focusing on risk-adjusted returns of ESG and traditional investments, they found that 33% of studies found a positive impact of ESG strategy on investment performance, 26% found it to be neutral, 28% had mixed results, and 14% found negative results. Then, covering 107 unique studies published since 2015, they concluded that ESG investing returns were generally indistinguishable from conventional investing returns. Notably, they highlighted that most studies showed ESG integration as a more profitable strategy than negative screening and divesting. Out of 17 studies, 33% found that ESG integration brought alpha, and 53% found neutral or mixed results.

All in all, the academic literature delivers no clear evidence for constant ESG investments' underperformance or outperformance. However, ESG investments typically do not underperform the traditional ones. Depending on the study period and the ESG strategy investigated, investments considering ESG-related issues might bring contradictory results. This is consistent with the underlying theory behind passive investing, as no strategy consistently beats the market over the long term.

2.6 Passive ESG Approach

While the majority of ESG investments apply active strategies, the ESG indexing industry has grown substantially in recent years. According to the Index Industry Association Benchmark Survey (2022), the number of ESG indexes increased by 55% in 2022, surpassing 50,000 globally, while the total number of indexes across all categories grew by only 4.43%. The development of ESG indices demonstrates a natural implication of the ESG industry growth and a response to customers' demand. The explosion of ESG data, ESG rankings and research has allowed the passive ESG approach to become widespread. Alternatively, passive ESG investing is known as optimized index investing or sustainable enhanced indexing (US SIF Foundation, 2020: 3).

The origins of ESG indexing date back to 1990 when the first ESG index was designed. Domini 400 Social Index was founded by KLD's Amy Domini, an American investment adviser. The index aimed to deliver exposure to the stocks of companies that KLD assessed positively from the perspective of their environmental, social, and corporate governance characteristics. The index has existed up to now as the MSCI KLD 400 Social Index. It consists of 400 companies from the universe of large, mid, and small-cap companies in the MSCI USA IMI Index. The construction of an index is based on both exclusions and additional considerations. It excludes securities of companies involved in nuclear power, tobacco, alcohol, gambling, military weapons, civilian firearms, GMOs, and adult entertainment, and then investigates ESG performance, sector alignment, and size representation to maintain similar sector weights as the MSCI USA IMI Index (MSCI KLD 400 Social Index (USD) Factsheet, 2023).

The first global ESG index, the Dow Jones Sustainability World Index (DJSI World), was launched in 1999. It tracks the top 10% of the largest 2,500 companies in the S&P Global BMI based on their performance in the Corporate Sustainability Assessment (CSA) (Dow Jones Sustainability World Index Factsheet, 2023). Over the years, the ESG indexing industry has started to offer more and more specialized and unique products. The first and nowadays most well-known clean energy index – the WilderHill Clean Energy Index – was introduced in 2004. The index is designed to represent the most prominent businesses that benefit from the transition toward the use of cleaner energy, zero-CO_2 renewables, and conservation. The composition and weighting of the index are based on securities' relevance regarding energy transition and technological advancements aiming to prevent pollution (WilderShares, 2023). The overview of the selected ESG equity indices is presented in Table 6.

Table 6. Examples of ESG equity indices by rating providers

ESG Rating Agency	Index Provider (Parent Company)	ESG Index Name	Year of Index Launch
Moody's ESG Sol. (Vigeo EIRIS)	Moody's	ESG Euronext-Vigeo (formerly ASPI indices)	2001
		Ethibel Sustainability Index	2002
Bloomberg	Bloomberg	Bloomberg MSCI Socially Responsible Indices	2014
		Bloomberg MSCI Sustainability Indices	2014
		Bloomberg MSCI ESG-Weighted Indices	2014
		Bloomberg MSCI Green Bond Indices	2014
		Bloomberg SASB ESG Corporate Indices	2019
		Bloomberg SASB ESG Equity Indices	2019
		Bloomberg Goldman Sachs Global Clean Energy Index	2021
		Bloomberg Rockefeller US All Cap Multi-Factor ESG Improvers Index	2021
		Bloomberg MSCI Socially Responsible Indices	2014
		Bloomberg MSCI Sustainability Indices	2014
		Bloomberg MSCI ESG-Weighted Indices	2014
		Bloomberg MSCI Green Bond Indices	2014
		Bloomberg SASB ESG Corporate Indices	2019
		Bloomberg SASB ESG Equity Indices	2019
		Bloomberg Goldman Sachs Global Clean Energy Index	2021
Refinitiv Eikon	LSEG	Refinitiv Eurozone ESG Select Index	2005
		Refinitiv/S-Network ESG Best Practice Indices	2017
		Refinitiv IX Global ESG High Dividend Low Volatility Equal Weighted Index	2004
		Refinitiv Global Resource Protection Select Index	2007
		The Thomson Reuters/Future Super Australia Fossil Free Index	2010
FTSE Russell	LSEG	FTSE4good index series	2001
		FTSE ESG index series	2015
Sustainalytics	Morningstar Group	Jantzi Social Index (JSI)	2000
		Global Sustainability Signatories Index (GSS)	2011
		STOXX Global ESG Leaders Indices	2018
MSCI	Morgan Stanley	MSCI KLD 400 Social Index (formerly KLD Domini 400 Social Index)	1999
		MSCI Fixed Income ESG index	2020
Calvert	Morgan Stanley	The Calvert US Large-Cap Core Responsible Index (formerly the Calvert Social Index)	2000
		Calvert International Responsible Index (formerly Calvert Developed Markets Ex U.S. Responsible Index)	2015
S&P Global ESG(Robeco SAM)	S&P	Dow Jones Sustainability Indices (DJSI)	1999
ISS ESG	Deutsche Börse	ISS ESG EVA Leaders Index series	2017
		ISS Governance Quality Score Index series	2021

Source: Antolín-López and Ortiz-de-Mandojana, 2023

Even though ESG indices are an extremely heterogeneous group, as every provider applies a different methodology, the general process of ESG index construction is most common (S&P Dow Jones Indices, 2023; MSCI, 2020; FTSE Russell, 2018; Weiner et al, 2020). Considering that ESG investing is about the integration of non-financial factors rather than a complete rejection of traditional metrics, most commonly, the basis for ESG indices is the traditional index. The parent index or the broad equity market defines the primary security universe. The next step in the security selection process includes screening, whose purpose is to exclude securities that are not aligned with the ESG strategy used by an index provider. Then, the remaining securities are ranked to achieve a specific ESG exposure or ESG-related theme. Typically, this involves assigning numerical scores to each company to select those with the highest scores or reject those with the poorest scores. Additionally, the ESG index might use additional criteria connected with the parent index, such as similar sector or regional weights. Once the securities are selected, their weights are established under the adopted method, like capitalization-weighting, equal-weighting, or others. Like traditional indices, ESG indices are rebalanced and reconstituted periodically, for instance, quarterly, semi-annually, or annually. Not only does it provide an opportunity to verify the index components concerning corporate actions, but it also checks the validity of ESG data.

Essentially, what distinguishes ESG indices from traditional indices is their reliance on ESG ratings and data specifically related to individual companies. These ratings assess firms' environmental, social, and governance performance, and are crucial for selecting and weighting constituents in ESG indices. The construction of ESG indices would not be possible without the whole ESG industry collecting information concerning companies' environmental, social, and governance issues, as well as rating providers who enable ESG index providers to select and rank securities in the portfolio. Thus, the objectivity of ESG scoring agencies is crucial for the ESG index quality. As indicated by Chatterji, Levine, and Toffel (2009), SRI raters do their best to convince clients that their methods are based on detailed analysis and high-quality data. Frequently, they apply numerous data sources from official government data to company documents, press releases, interviews, and surveys. Moreover, they use multiple research methods and techniques to protect against methodological subjectivity.

However, one major problem with ESG ratings is that they are not consistent with each other. The reason for this is the use of different methods and variables to measure the same issues. Delmas et al. (2013) pointed out that depending on the ESG rating provider, environmental performance might be measured with indicators of a firm’s environmental processes or with metrics of the company’s environmental outcomes, which causes significant discrepancies. Chatterji et al. (2016) examined the convergence of SRI raters applying the ratings of companies from six leading social raters: KLD, Asset4, Innovest, DJSI, FTSE4Good, and Calvert. They documented the lack of agreement across social ratings between different SRI ratings providers, even after the adjustment for explicit differences in the definition of CSR. Then, Berg et al. (2022) extended the research by decomposing the divergence into scope, weight, and measurement, showing that this is the measure component that is mostly responsible for the ESG rating discrepancies. This means that, essentially, different rating providers do not agree on which categories are most important in the ESG evaluation. The authors showed that the correlations between ratings encompassing the complete set of 709 underlying ESG indicators provided by KLD, Sustainalytics, Moody's ESG, S&P Global, Refinitiv, and MSCI ranged from 0.38 to 0.71.

The consequences of disagreement between different ESG rating providers are also relevant for ESG index providers, as it affects the reputation and trust of the whole ESG industry. From another perspective, different approaches to ESG scoring represent a possibility for investors to go for the rating methodology that best suits their investment strategy. However, as long as the ESG industry does not exhibit a sufficient level of transparency and a strict division framework, it just causes chaos and misunderstanding.

To some extent, ESG and passive investing might seem to be contrary concepts. As far as returns are concerned, Cornell and Damodaran (2020) indicated that there is no evidence that active ESG investing performs better than the passive ESG approach, however, the contradictions emerge in the sheer nature of ESG investing. Notably, one may argue that passive ESG investing is not passive because an ESG strategy implies a shift away from traditional capitalization-weighted indices (Hallez, 2022). The question is whether passive ESG investing, which involves active tilting of the portfolio away from certain sectors or companies, introducing a level of active management, still might be called passive investing. While the essential goal of passive investment is to achieve market-like returns at a low cost, ESG investing incorporates environmental, social, and governance factors into the investment process, seeking to align with certain non-financial considerations. It may lead to some deviations from the real market portfolio composition, resulting from exclusions and altered weightings.

What is more, passive ESG funds are criticised due to their reliance on conventional indices, which are unsustainable by nature (US SIF Foundation, 2020). Rompotis (2023) conducted a study on the correlation between the returns of ESG index funds traded in the U.S. and the return of the S&P 500 Index, showing a mean correlation equal to as much as 0.90. Out of 27 index funds, 18 funds exhibited correlations with the S&P 500 Index of 0.90 or higher, and the lowest correlation in the sample was 0.65. Focusing on holdings of the investigated ESG funds, Rompotis (2023) proved that 26% of the funds invested at least 5% of their assets in stocks with poor ESG profiles. The Author suggests that ESG funds commonly use some kind of greenwashing practices.

Another point of controversy is the cost of adopting a passive ESG strategy. Baker et al. (2022) found that between May 2019 and March 2022, non-ESG index mutual funds were more likely than ESG-oriented funds to have very low expense ratios (below 10 basis points), but the median expense ratio was higher for non-ESG funds than for ESG funds. Nevertheless, after controlling for fund characteristics such as ESG mandate, fund age, sponsor fixed effects, and market-by-month fixed effects, it was found that ESG funds had expense ratios that were 4.6 basis points higher per annum than non-ESG funds.

On the other hand, ESMA Annual Statistical Report (2022) *Performance and Costs of EU Retail Investment Products* indicated that in 2020, ESG funds in Europe were less costly than their non-ESG peers, even after controlling for the age and the size of the funds. However, having distinguished between ETFs and non-ETFs, they found significant differences. The average ongoing costs in ESG equity UCITS funds (excluding ETFs) were 1.5% compared with 1.8% for non-ESG peers due to TER differentials. Conversely, in the ETF segment, ESG equity ETF UCITS total costs accounted for 0.8% versus 0.6% for non-ESG equity ETFs.

Having in mind the additional costs associated with possessing ESG data, conducting research, and implementing ESG strategy, it seems counterintuitive that ESG funds might offer lower fees than non-ESG funds. However, as suggested in the EFAMA (2021) report, the reasons for the relative cheapness of ESG funds may derive from the growing competition in the ESG industry. Over the years, ESG funds have substantially lowered their fees to stay competitive in comparison to non-ESG funds and within the ESG industry.

2.7 Summary

The considerations on ESG investing presented in this chapter aim to explain how the ESG approach is different from other investing strategies and how it works. In particular, the chapter focused on presenting the key characteristics of ESG investing, its strategies, dimensions, and performance. The chapter did not comprehensively analyse ESG investing, but rather signals some most important problems, especially for passive ESG products.

The review of studies and reflections on the current methods of ESG investing incorporation allowed us to prove that ESG investing, as any innovation, has much to do with regard to systematization and transparency. As a modern and future-oriented approach, ESG investing has great potential for growth. However, what is holding the industry back is insufficient systematisation. Essentially, ESG investing struggles to broaden the scope of traditional investment analysis by incorporating into the decision-making process environmental, social and governance issues. It does not diminish the importance of financial factors but makes the analysis more complete by analysing non-financial issues thoroughly. In practice, the implementation of ESG strategy takes various forms from negative screening to shareholder engagement. Then, the divergences in understanding and implementing the ESG strategies open up opportunities to abuse in terms of greenwashing. Numerous ESG initiatives and standards existence seem to be insufficient as ESG investing is commonly criticised for its inconsistencies, lack of standards, and comparability.

Finally, the chapter explored the passive ESG approach as the base for passive ESG equity ETFs, constituting the primary interest of the thesis. It was crucial to get accustomed to the strategy before proceeding to the research on passive ESG ETFs' tracking ability. All in all, as the passive ESG approach principally relies on the quality of quantitative ESG metrics, the lack of consistency among various ESG data and ratings providers constitutes a significant obstacle to industry development. Essentially, it breaches investors' trust in the industry. This is why further clear ESG regulations and standards are needed.

CHAPTER 3 Benchmark Replication in ETFs

3. 1 Introduction

Investment funds are institutional units, excluding pension funds, that consolidate investor funds to acquire financial assets (IMF, 2004: 161). Investment funds are managed by professionals who select and oversee the portfolio, giving investors access to diversification without requiring detailed knowledge of individual assets. Within the broad category of investment funds are mutual funds, exchange-traded funds (ETFs), hedge funds, and private equity funds. ETFs are a form of investment fund that developed from mutual funds and allow investors similar diversification and pooled investment as mutual funds, but with the added advantage of intraday trading. The primary objective of passive ETFs is to create an investment portfolio that replicates the performance of a given index as closely as possible (Bodie et al., 2014: 95).

Initially, all funds were actively managed, but with the development of the theory of finance, the passively managed funds industry began to grow, first index funds and then passive ETFs. The creation of the First Index Investment Trust in the USA - the Vanguard 500 (VFINX) in 1976, followed by further index funds opened in the 1980s, has started the expansion of the passive investing approach. However, despite a strong theoretical foundation and the justification of this approach to investing proven by numerous studies, it has not enjoyed significant investor interest for many years. Passive investing became widespread after the invention of a new investment product that combined passive portfolio management with the possibility to trade its units on the exchange – exchange-traded fund (ETF). The first fund of this kind, known as the Toronto 35 Index Participation Units (TIPs), was launched on the Toronto Stock Exchange in 1990 (Toronto Stock Exchange, 2025). In the United States, the first ETF – Standard & Poor's Depositary Receipts (now SPDR S&P 500 ETF Trust) – was introduced in 1993 by State Street Global Advisors (SSGA, 2025).

Since the 1990s, the ETF industry has grown enormously, not only in terms of capitalisation and the number of such funds worldwide but also in terms of the plethora of investment opportunities offered by these products. Initially focused on equity indices, the ETF industry has begun to include other asset classes such as fixed-income securities, REITs, and commodities. Furthermore, more and more sophisticated products have started to arise, such as smart-beta funds, factor-based ETFs, sustainable ETFs, and actively managed ETFs. As this dissertation is devoted to a special category of funds - passive ESG equity ETFs

- before proceeding to the analysis of this one fund type, it is necessary to outline the broad ETF industry. Thus, this chapter will provide the most significant information about ETFs, the specifics of their functioning, pros and cons, tracking ability issues, and their determinants.

3. 2 Essentials of Exchange-Traded Funds

An ETF is a pooled investment vehicle whose shares (units) can be traded throughout the day on an exchange at a price set by the market. Like a mutual fund, an ETF offers investors a share in a basket of assets under the fund's management in proportion to the number of units purchased. Although ETFs may resemble mutual funds in many aspects, they have some underlying differences presented in Table 7.

Table 7. Key differences between ETFs and mutual funds

Feature	Mutual funds	ETFs
Buying and selling	Through the mutual fund company or broker, at the NAV price after the market closes	On exchange throughout the market trading day at real-time price
Transparency	Quarterly information about holdings on at least a 30-day lag	daily information about holdings
Minimum investment	May exist, usually from USD 500 to USD 5000	One share
Management	Typically active	Typically passive
Expense ratio	Generally higher	Generally lower
Trading costs	Costs of inflows and outflows are shared by all holders, which reduces the NAV	ETF buyer/seller bears trading costs
Tax efficiency	Higher tax resulting from frequent transactions	Considered tax-efficient

Source: Own elaboration based on WisdomTree (2021)

Most commonly ETFs are structured as open-end investment funds (OEFs), which means that they are governed by the same regulatory documents as mutual funds. In the United States, ETFs are typically structured as OEFs and the principal legislation governing their operation is the Investment Company Act (1940). In the European Union, in turn, exchange-traded funds are commonly structured as UCITS funds (Undertakings for Collective Investments in Transferable Securities) and regulated by the UCITS Directive (2009). Exchange-traded funds structured as open-end fund are deeply restricted with regard to their investment policy. This involves high diversification requirements and limited access to alternative financial instruments and assets such as commodities and currencies. The ability to reinvest dividends immediately, the use of derivatives, portfolio sampling and securities lending can be pointed out as advantages of this ETF legal structure. Alternatively, exchange-traded funds in the USA can be structured as unit investment trusts (UITs), which combine features of mutual funds and closed-end funds (CEFs). Because such funds trade in a fixed basket of assets up to the termination date, and cannot lend securities, nor use sampling they are less flexible than ETFs structured as open-ended funds (Miziołek et al., 2020: 68–69).

Originally, ETFs were created with one primary purpose - to track the performance and risk of a selected benchmark, which is most often a financial market index. Over more than 30 years in the business, ETFs have gradually evolved and broadened the spectrum of investment opportunities for their clients. Nowadays, ETFs can be structured to track anything from the price of a single security or commodity to a large and diverse collection of securities and even specific investment strategies. Passively managed ETFs have continuously comprised the majority in the market, however, there has also been growth in the actively managed ETF industry. According to ETFGI in 2023 active ETFs accounted for 5.5% of the global ETF market (ETFGI, 2023).

Even though it might seem that a passive ETF strategy is simple and undemanding, in reality managing a passive fund is not at all easier than managing an active fund. This is because indices themselves are not investable and are merely synthetic statistical indicators. Deviations from the performance of the index tracked by the ETF are the result of numerous difficulties, beginning with the high complexity of the index, changes in its composition and weights, to unforeseen market situations. Therefore, a passive ETF is not at all passive in its portfolio management activities, it just does not actively attempt to beat the performance of the broad market (Miziołek et al., 2020: 72–73).

3. 2. 1 Operation of Exchange-Traded Funds

The process of launching an ETF on the market is multi-stage and demands numerous requirements to be met. As a first step, the fund sponsor has to specify several issues regarding both the instrument and the way it will function. Besides determining the fundamental considerations for each open-ended fund, such as the objective, investment strategy, asset and risk management methods, domicile, jurisdiction, taxation, or accounting, an ETF must provide additional specific information. This includes matters such as index licensing, listing platforms, agreements with market makers, authorised participants and liquidity providers, as well as NAV and iNAV calculation agents and security borrowers (Lewellyn, 2016). Then, an ETF sponsor files a request to the financial supervisory body to register an ETF and subsequently, to introduce its shares on the market.

A unique feature of ETFs is that their units are traded on two markets: the primary market and the secondary market. The primary market for exchange-traded funds represents something substantially different from the primary market for other instruments such as stocks. While the primary market for equities involves initial public offerings (IPOs) and capital flows from investors to stock issuers, the purpose of the primary market for ETFs is the creation and redemption of its shares. The process of creating and redeeming ETF shares takes place between the fund and a large financial institution or market maker (an authorised participant). Unlike stocks, whose listing on the primary market occurs only once (at the time of the initial public offering), the creation of ETF units is a continuous process and takes place even on a daily basis. On the primary market for exchange-traded funds, their units are traded at the fund's net asset value (NAV) at the end of the day usually in very large holdings called *creation units*, which comprise several or even tens of thousands of ETF units (usually from 25,000 to 200,000 shares) (Antoniewicz and Heinrichs, 2014).

A typical ETF unit creation process comprises an Authorised Participant (AP) collecting the underlying securities of the fund in the appropriate weightings and delivering them to the fund issuer (see Figure 11). Then, the ETF issuer uses a continuous issuance function to provide the AP with new shares that constitute the same value as the delivered securities. The process increases the number of outstanding ETF units. It is mutually beneficial for both transaction parties, as the ETF issuer gains securities so it can enter the market and an AP receives ETF units, which can then be sold profitably on the open market (Abner, 2010: 46). This method of ETF shares creation in which an AP delivers a basket of securities to the fund sponsor in exchange for ETF units is called *in-kind creation* or *in-specie creation*. This mechanism is used most often by ETFs applying physical replication. ETFs using synthetic replication are more likely to use the cash creation method to create new ETF shares. This method is based on the transfer of cash by an AP to the fund in exchange for ETF shares. The amount transferred to the fund is equal to the value of the underlying securities comprising the fund's portfolio, usually purchased through program trading (BlackRock, 2010: 5)

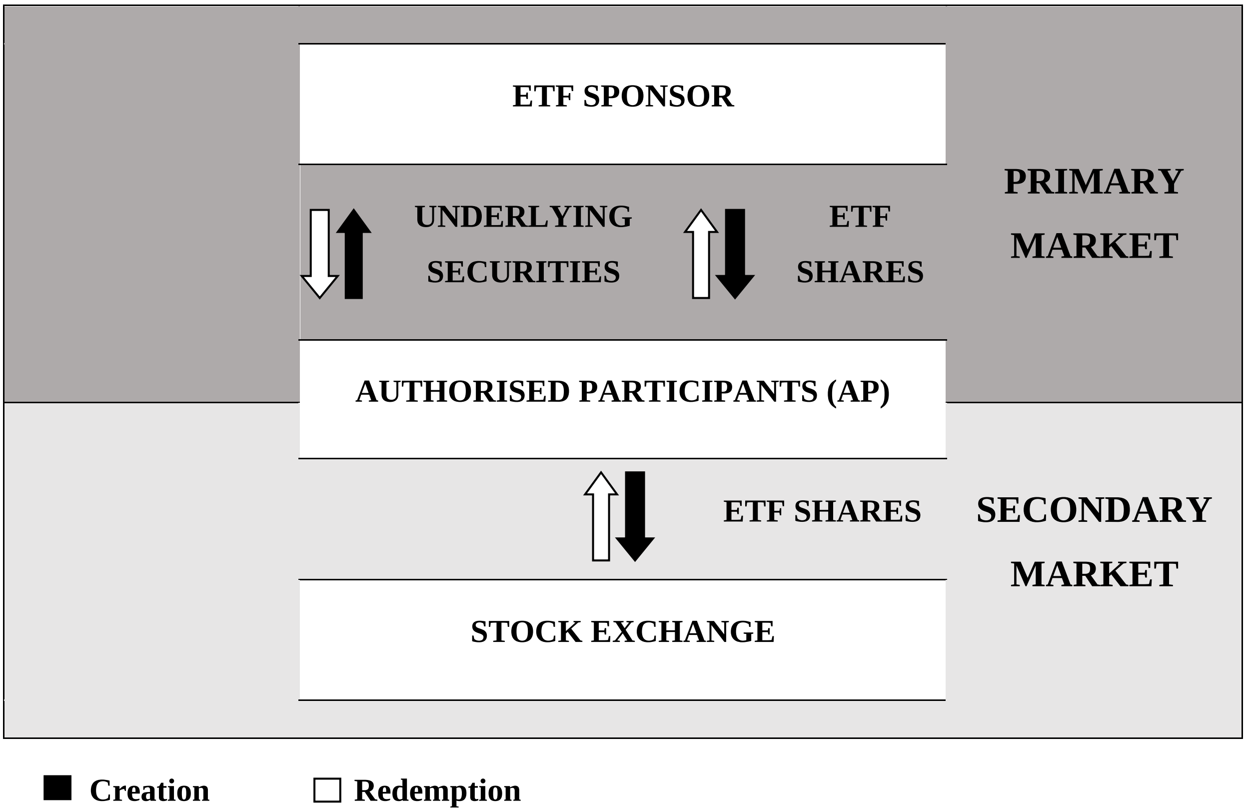

Figure 11. Operation of Exchange-Traded Funds

(Source: State Street Global Advisors)

The creation and redemption of ETF units are ongoing processes performed by authorized participants (APs) to maintain the market price of ETF shares close to their net asset value (NAV) despite temporary deviations caused by supply and demand pressures in the secondary market. When ETF shares are traded at a premium, which means that their market price is higher than their NAV, authorized participants (APs) step in to profit from the price difference. They buy the same securities that the ETF holds and exchange them for new ETF shares from the issuer. Then, they sell those shares on the stock exchange at a higher price, earning a profit. When ETF shares are traded at a discount, which means that their market price is lower than their NAVs, APs buy them on the market and return them to the issuer in exchange for the underlying securities. This reduces the number of ETF shares and helps bring the price back in line with the fund's net asset value. This process is known as arbitrage and ensures effective valuation of ETF units (Madhavan, 2016).

Trading ETF units on the secondary market differs from transactions on the primary market. While in the primary market, the applicable transaction price is the NAV, in the secondary market, transactions are executed at the ETF share market price, which may at times, deviate from fair value (as long as APs do not arbitrage). Moreover, due to the very high minimum transaction value on the primary market, hardly any individual investors trade on it. In the secondary market, the minimum transaction size is one ETF unit, so the participants in this market include not only institutional investors but also minor individual investors. Finally, as the primary market is characterised by higher liquidity, transactions performed there have minimal impact on the price of ETF units. It is different in the exchange, especially in immature small exchanges, where the price can fall or rise significantly even as a result of a single transaction (BlackRock, 2010: 4). Importantly, in typical market conditions, the majority of ETFs do not have any activity on the primary market, so they do not create or redeem shares. As long as an ETF has sufficient liquidity, its activity is focused on the secondary market to trade its shares (Investment Company Institute, 2014).

Although trading of ETF units and stocks is similar in many basic aspects, one should realize that there are also important differences between these trades. Common features include trading on the same markets in a continuous trading system, publication of instrument prices and current bid and ask prices in real-time, as well as their price sensitivity to market conditions. On the contrary, there are significant differences between trading ETF units and stocks. Multi-listing is quite typical for exchange-traded funds, especially on European markets, while individual stocks are very rarely listed on two or more exchanges simultaneously. Finally, there are divergences in the publication of the instruments' value. While the valuation of stocks is limited to market value, the valuation of ETF shares additionally involves the periodic publication of their theoretical estimated value – indicative NAV (iNAV) (Gastineau, 2010: 197–201).

3. 2. 2 Methods of Index Replication

The aim of any passive fund, such as an index fund or passive ETF, is to replicate the performance of the tracked index as closely as possible. Therefore, this part of the paper presents the principal methods by which passive fund managers try to achieve this goal. The primary division of index replication methods distinguishes between physical replication and synthetic replication. While physical replication involves the purchase of securities for a portfolio designed to track the index, synthetic replication uses derivatives to do so.

The first and basic physical index replication method is full replication. This method is based on purchasing into the fund's investment portfolio all of the securities comprising the underlying index. This means that the portfolio of such a fund consists of the same instruments with identical or substantially similar weightings as the index. The advantages of the full replication method certainly include transparency and a relatively low tracking error. The participants in funds applying this method are provided with real-time information on the composition of the portfolio, which most often corresponds to the index portfolio. As the number of securities in the portfolio increases, the tracking error decreases because the passive portfolio replicates the index more and more accurately. However, when the manager adds illiquid and non-significant instruments to the portfolio, it can significantly increase trading costs and make it harder for the fund to match the index's performance. Therefore, the full replication method is used primarily to track the performance of equity indices comprising the most liquid and largest companies, for instance, the S&P 500, or the FTSE 100. Used to track the performance of small-company indices, especially in emerging markets, this method will prove ineffective because of the excessive costs (CFA Institute, 2019: 119).

As holding in the portfolio all the constituent securities from the index often proves to be impractical, portfolio managers of passive funds often decide to use the stratified sampling method to replicate the index. The stratification process encompasses arranging the whole population into exclusive and exhaustive subgroups based on several different criteria. The referencing points may concern the company's market capitalization, sector affiliation, or the value of market ratios, as well as many others. The benefit of stratifying along multiple dimensions is the improvement of the tracking ability. Once each company is allocated to a category, the weighting of each sub-group is determined. Finally, when selecting companies for the portfolio tracking the performance of the index, the fund manager purchases only a portion of the securities in each category, most often excluding the illiquid ones. Simultaneously, the manager maintains the weights of the particular categories so that the sum of the weights of the instruments from each category in the created portfolio corresponds to the sum of the weights of the given category in the index. The most significant advantage of the stratified sampling method is the reduction of costs associated with acquiring an excessive number of instruments while maintaining the characteristics of the replicated index. For this reason, the method is particularly suitable for the replication of broad market equity indices with many components, such as the S&P Global Broad Market Index, which consists of more than 11,000 constituents (CFA Institute, 2019: 121).

Next, there is an optimisation technique which is undoubtedly the most advanced physical index replication method. The underlying purpose of this technique is to maximize desirable characteristics or minimise undesirable characteristics. For example, as a desirable feature, one may consider minimising index tracking error or limiting the number of securities in a portfolio. To select securities for the fund's portfolio, it uses sophisticated mathematical models. Among the predominant models used in the optimisation technique are multifactor risk models, which measure the risk exposure of an index and individual securities. In comparison to the stratified sampling method, an important advantage of this technique is that it considers not only the individual factors that explain the behaviour of security prices, but also their covariance (CFA Institute, 2019: 122). On the other hand, as the method is based on historical data, the possibility of inadequate risk modelling and overfitting the data is indicated as its drawbacks. Roll (1992) and Jorion (2004) pointed out that applying the optimisation technique to minimize tracking error can lead to portfolios that are mean-variance inefficient versus the benchmark.

One advantage of the physical replication method used by an ETF is the possibility of gaining additional income through securities lending agreements. Securities lending depicts the temporal lending of securities to another entity for a certain fee. This possibility is essential from the perspective of the fund's efficiency, as it improves the investment result and reduces the tracking error, sometimes allowing an ETF to generate even a better result than the underlying benchmark. On the other hand, securities lending may also constitute an additional counterparty risk (associated with the possibility that the other party will fail to honour the obligation). However, the risk might be substantially reduced by imposing mandatory collaterals (Morningstar, 2012: 4–5).

The second type of replication strategy used by ETFs is synthetic replication. Synthetic replication, also known as swap-based replication, involves the use of a derivative – usually a swap contract, between the investment bank and the ETF provider. The method does not assume an investment in the underlying markets but rather maps them. To replicate the performance of an index through a swap contract, the counterparty agrees to pay or receive any difference between the return on the index and the return on the basket of securities held. Back in 2010, 46% of assets in equity ETFs and 35% of assets in fixed income ETFs in the European market were held in synthetic funds. However, by the end of 2020, this market share had shrunk to just 17% and 5% respectively. On the other hand, synthetic replication structures are becoming increasingly popular for ETFs with exposure to mainstream US equities - from 2017 to 2020, the market share of synthetic ETFs tracking the S&P 500 and MSCI USA indices increased by 10 percentage points to around 30% (Morningstar, 2021).

Two main approaches are applied by synthetic ETFs: the unfunded swap model and the funded swap model. The unfunded model involves using cash raised from ETF investors to purchase a substitute or reference basket of securities from a swap counterparty. This basket is designed to deliver the return of the reference index. Although the substitute basket, as the name suggests, is only a substitute and does not contain constituent instruments from the index benchmark, most often it has a high degree of correlation with the index portfolio (Morningstar, 2012: 7). Most often, such a substitute portfolio contains instruments of the same type as the underlying benchmark (e.g. equities or bonds) distinguished by high liquidity and low risk.

Then, the funded swap model involves the transfer of investors' money by the fund issuer to the counterparty in exchange for the index return (reduced by swap fees) plus the value of the capital in the future. The counterparty uses this cash to purchase a basket of securities and deposits this substitute basket as collateral in a separate account, which is pledged to the fund. This means that technically this basket of assets is the property of the fund and, as such, provides protection to investors in the event of the counterparty's bankruptcy (BlackRock, 2010: 3–4).

One undoubted advantage of the synthetic replication structure applied by ETFs is the lower tracking error than funds using physical replication. Predominantly, this is because the return on the underlying index is guaranteed by the swap counterparty (Morningstar, 2021). Furthermore, synthetic index replication is often a more efficient and less costly solution for ETFs with exposure to less liquid or niche markets. On the other hand, an inherent element of synthetic replication is the counterparty risk (Vanguard, 2021: 12). According to Morningstar (2021), it is counterparty risk that is the major reason for the declining popularity of synthetic ETFs in Europe.

All in all, both physical and synthetic replication have their pros and cons. While the physical replication method is valued for its simplicity and transparency, synthetic replication, although more complicated, provides greater tracking quality and often proves to be a more cost-effective solution.

3. 2. 3 Valuation Mechanisms and Price Dynamics

Another fundamental issue related to the operation of ETFs is their valuation. The primary measure of ETF unit value is the net asset value (NAV), which reflects the market value of all the fund assets less the value of its liabilities. It is based on the most recent closing prices of the constituent securities in the fund and the total amount of cash on a given day. To calculate the NAV per unit, the total NAV value should be divided by the number of ETF shares (Abner, 2011).

While calculating the NAV value for domestic ETFs does not seem difficult, it is no longer so obvious for ETFs tracking international market indices. This is mainly due to the time zone and currency differences that occur. If a fund's portfolio includes securities from different stock exchanges, it is often the case that the fund's units are still listed on a given day, and the trade of underlying securities noted on another exchange has already been finished since the exchange is closed. This leads to a situation where the valuation of ETF units continues to change, while the price of the underlying securities remains unchanged. As the NAV is calculated based on the most recent closing prices of the constituent securities, there are inexorably divergences between the NAV and the market valuation of the ETF units. Inaccuracies in the NAV calculation also occur when the constituent securities are denominated in different currencies. To convert the value of securities into home or other base currency, most ETF providers use the WM/Reuters 4 p.m. London fix. As a result, the NAV does not reflect the actual value of the currency at the closing of the underlying securities, but at a single artificially selected point in time (Jane Street, 2018).

Another method that eliminates some of the problems mentioned earlier is intraday net asset value (iNAV), also known as indicative NAV. This measure provides nearly real-time information on the value of ETF assets during the trading day (typically the iNAV value is published every 15 seconds). The iNAV value is calculated based on the most recent market prices of all the constituents in the basket. It constitutes the current aggregated value of all the securities in the fund creation unit (calculation basket) adjusted by the cash components (Abner, 2011).

The iNAV value is calculated independently of the current price of the ETF units on the secondary market. However, it should be noted that for domestic ETFs, these two independently generated values should be consistent with each other. As with NAV, inconsistencies arise when ETF units are traded in different time zones than the underlying securities.

Turning to a discussion of the ETF market valuation, it should be noted that the price reflects the current supply and demand for ETF shares on the secondary market. It is the highest price at which a buyer is willing and able to purchase an ETF unit, and simultaneously the lowest price to which a seller can agree. The market price often deviates from the NAV value of ETF units. Divergences between market valuation and the fair value of ETF units arise when there are short-term imbalances in supply and demand. One example is when a large order for ETF units arrives at the very end of the trading day, and it is too late for AP to arbitrage. Most often, however, premiums and discounts apply to ETFs tracking illiquid securities, especially foreign assets listed in different time zones than the ETF.

Numerous scientific studies have confirmed that U.S. ETF prices are very close to NAVs, and this is the redemption feature that ensures ETF pricing efficiency (Ackert and Tian, 2000; Elton et al., 2002). The researchers have also pointed out that although domestic ETFs are usually adequately priced, mispricing for international ETFs is more common (Jares and Lavin, 2004; Engle and Sarkar, 2006). According to Ackert and Tian (2008), time zone differences do not fully explain the occurrence of the fund premium of international ETFs. They indicated the relation between ETF mispricing and momentum, illiquidity, and size effects, which may represent capital controls and other market limitations.

3. 2. 4 ETF Liquidity

As one feature of exchange-traded funds is that their units can be traded on stock exchanges (or over-the-counter secondary markets), liquidity is an important issue concerning their operation. First of all, it is necessary to provide a clarification of what liquidity means since it is not a one-dimensional phenomenon. Von Wyss (2004) enumerated the following dimensions of liquidity as the most commonly mentioned in the literature:

— Trading time – the ability to execute transactions at the current price without delay. This can be measured as the waiting time between transactions or the number of trades per unit of time.

— Tightness – it is the ability to buy and sell securities at roughly the same price at the same time. This reflects the cost related to executing the transaction and can be measured in different spread versions.

— Depth – it is the ability to buy or sell a certain amount of an asset without affecting the quoted price. This dimension reflects the volume at the best bid and ask prices and can be measured by the order ratio, the trading volume, or the flow ratio.

— Resiliency – the ability to buy or sell a certain amount of an asset with little influence on the quoted price. It reflects the elasticity of supply and demand and can be measured by the intraday returns, the variance ratio, or the liquidity ratio.

Ultimately, liquidity pertains to the ability of willing buyers and sellers to exchange assets at mutually agreeable prices, which is fundamentally dependent on specific technical conditions of a market. As ETFs buy and sell the demand for exposure to particular asset classes, market makers can net off buying and selling demand more efficiently than single securities. Consequently, an exchange of ETF shares between an institutional investor and the other side is less risky for the latter, and thus, market makers may make the market tighter (lower spreads).

This is why ETFs commonly achieve a higher degree of liquidity than the underlying holdings and can be traded at a lower expense (Golub et al., 2013: 9-10).

While for equities, average daily volume (ADV) is the primary indicator of liquidity, in the case of ETFs, ADV reflects the liquidity of the fund only to a certain extent. This is because, unlike equities, which have a fixed number of shares (until the next offering), the number of ETF shares changes frequently as a result of the creation and redemption processes. The liquidity of exchange-traded funds is multi-layered. This means that ETFs have different levels of liquidity, which allow investors to trade ETFs in amounts that can significantly exceed the fund's ADV without noticeable impact on the fund's price. There are three levels of ETF liquidity (Golub et al., 2013: 10):

— displayed liquidity – liquidity displayed “on screen” in the secondary market (market makers do not display publicly all the information);

— reserve liquidity – non-displayed secondary market liquidity (the information about this level may be sourced through relationships with market makers);

— primary market liquidity – “true” liquidity (underlying basket).

In the primary market, the liquidity of ETF units is determined foremost by the liquidity of ETF underlying securities. When the constituent elements of an ETF basket are sufficiently liquid, usually the AP has no problems with providing the primary ETF market liquidity by carrying out the creation and redemption processes. In the secondary market, apart from the liquidity of constituent instruments comprising an ETF basket, fund liquidity is also influenced by the activities of market makers and the spread between the bid and ask offers.

Market makers are responsible for maintaining sufficient liquidity in the ETF market by ensuring the continuous tradability of shares. They are broker-dealers who constantly provide buy and sell quotes of ETFs on the terms and conditions set out in the agreement with the exchange. Exchange-traded funds may have one or more market makers. However, it should be emphasised that it is advantageous for a fund to have numerous market makers to ensure high price competitiveness. Not only does it lead to lower spreads and higher liquidity, but also, consequently, lower costs (Golub et al., 2013: 11). This shows that market makers play a special role in the efficient operation of ETFs. Their role is particularly important for ETFs that do not attract a lot of investor interest, e.g., due to unusual or exotic market exposure.

Like any other security, an ETF share has two prices: a bid (the price at which someone is willing to buy a share) and an ask (the price at which someone is willing to sell a share). Obviously, sellers want to get the highest possible price for selling the security, while buyers want to acquire the instrument as cheaply as possible. The difference between these two prices - the spread - occurs precisely because of the negotiations between sellers and buyers. An ETF's share can only change hands when the buyer agrees with the seller, so the bid and ask prices will be equal. The size of the spread is expressed on the exchange in basis points and represents the cost to the investor of opening or closing a position in the market. Most often, the maximum acceptable spread size is specified in the contract between the market maker and the exchange. The size of the spread depends on several factors, including (Global Asset Management Inc., 2017):

- spreads on the underlying securities – if the underlying securities are not liquid, an ETF could also have liquidity problems, and consequently – a considerable spread.
- cost of assembling and trading the basket of securities – typically an ETF that invests in foreign securities, has higher costs, which are reflected in the spread size.
- trading volumes – if an investor makes a very large purchase of ETF shares (exceeding the amount of inventory the market maker has for sale), the market maker may need to create more ETF units by buying large amounts of the underlying securities. This often requires having to pay higher asking prices to fill the orders, and in turn leads to a higher ETF spread.
- market risks – during volatile times, spreads tend to be higher.

In general, as demonstrated above, an ETF's liquidity depends on various factors. However, due to market makers' activities, exchange-traded funds usually exhibit very high liquidity. This means that in the case of insufficient liquidity of ETF shares on the secondary market, their liquidity is maintained by APs, so free rotation of ETF shares is guaranteed.

3. 3 Advantages of Exchange-Traded Funds

Having discussed how ETFs operate, it is time to move on to their predominant characteristics, representing also their advantages, which have strongly contributed to the growing demand for these financial instruments. This section presents the key ETFs' benefits and confronts them with reality and the latest research in this regard.

3. 3. 1 Portfolio Diversification

Diversification of an investment portfolio refers to a wide range of activities aimed at dispersing (and thus reducing) the risk of the entire portfolio through an appropriate selection of its components. While self-diversification of a portfolio by the individual investor was a labourintensive and costly task, with the emergence of mutual funds, especially low-cost index funds, every investor has access to a diversified investment portfolio at their fingertips.

There are two ways to diversify a portfolio: vertically and horizontally. Vertical diversification is based on reducing risk by investing in different asset classes. As noted by O'Sullivan and Sheffrin (2003: 273), if the prices of different assets do not move in perfect synchrony, a diversified portfolio will have less variance than the weighted average variance of its components, and often less volatility than the least volatile of its components.

The second method of portfolio diversification – horizontal diversification, refers to investing in only one asset class, but with dispersed characteristics. This diversification can be based on geographic dispersion, company size, sector affiliation or the company's investment style. The diversification benefits have already been demonstrated by Elton and Gruber (1977). While the expected variance of a portfolio consisting of a single security was 46,811, for a 20-component portfolio it was already 9,036 and for a 1,000-component portfolio 7,097 (Elton and Gruber 1977: 425). In this regard, exchange-traded funds, which comprise a basket of securities, offer investors a low-cost diversified portfolio.

As the level of diversification depends primarily on the correlation between the individual components of a portfolio, it would appear relevant to recall some studies in this regard. In an age of globalisation and increasing interdependencies between countries and markets, international diversification is of particular importance. Following the publication *Internationally diversified portfolios: Welfare gains and capital flocks* by Grubel (1968) providing the extension of portfolio analysis to the international market, academic interest in this matter increased.

Coeurdacier and Guibaud (2011) indicated that investors actively rebalance their portfolios to achieve more exposure to countries providing superior diversification effects. Then, Forbes and Rigobon (2002), Ratanapakorn and Sharma (2002), Leong and Felmingham (2003), and Dalkir (2009) evidenced that correlations between markets increase during volatile times. Accordingly, Beine, Cosma, and Vermeulen (2010) argued that trade and financial integration have increased the probability that international equity markets will crash jointly. Finally, Bekaert et al. (2009) and Christoffersen et al. (2012) remarked that correlations between international markets have increased notably in both developed and emerging markets. This means that with the increase in the intensity of international cooperation, investors' diversification opportunities have decreased.

There were also studies focusing specifically on ETFs' diversification benefits. Phengpis and Swanson (2009) argued that country ETFs constitute a good international diversification opportunity for U.S. investors since they are exposed to the movements of their underlying countries' indices more than is the case in the U.S. market. On the other hand, Tse and Martinez (2007) analysed the price discovery process and information transmission of twenty-four international iShares funds and found limited diversification benefits of ETFs. Filippou et al. (2022) noted that it is the ETF arbitrage mechanism that makes global shocks propagate to local economies. This, in turn, leads to increased return correlation with the U.S. market, limiting the benefits of international diversification. All in all, recent studies suggest that the diversification benefits of ETFs have decreased over time. Nevertheless, ETFs still offer a cost-effective way to achieve a certain degree of diversification.

3. 3. 2 Cost Efficiency and Trading Flexibility

Back in the chapter on passive portfolio management, it was pointed out that costs in passively managed funds are significantly lower than in actively managed funds. This is also applicable for exchange-traded funds.

Chen et al. (2022) examined how the expenses of 265 passive mutual funds deteriorate their performance. They juxtaposed ETFs and index funds tracking the same indices and managed by the same financial services firms as the Vanguard S&P 500 Index Fund and the Vanguard S&P 500 ETF. The study showed that the average annual expense ratios for passive mutual funds were more than double that of ETFs' average annual expense. The average cost of index funds was 0.45%, while ETFs were cheaper, with an average cost of 0.21%. The Authors pointed out that the underperformance of passive mutual funds results from investor service costs and the cash drag. They said that fees related to administration, record keeping, and regulations mandated by federal agencies were responsible for about 78% of the additional costs. The rest of the underperformance was caused by the cash drag, which relies on holding an excessive percentage of the funds’ assets in cash, which reduces the potential for gains.

Focusing on European funds, the ESMA Market Report on Costs and Performance of EU Retail Investment Products (2023) showed that over 10-year periods, ETFs were on average less expensive than both active funds and passive funds that were not ETFs (see Table 8). For example, between 2011 and 2020, the average ongoing cost for ETFs was 0.25%, compared to 0.46% for non-ETF passive funds and 1.50% for active funds.

Table 8. Average annual ongoing costs of EU UCITS products (in %)

Years	Active funds	Passive funds (non-ETFs)	ETFs
2009–2018	1.50	0.50	0.30
2010–2019	1.50	0.40	0.30
2011–2020	1.50	0.46	0.25
2012–2021	1.47	0.48	0.31

Note: UCITS = Undertakings for Collective Investment in Transferable Securities

Source: Own elaboration based on ESMA (2023)

Next, one should consider what is the reason for the lower costs of ETFs, even compared to passive index funds. Undoubtedly, the key advantage of ETFs is their unique feature connected to the creation-redemption mechanism. While passive index funds impose redemption fees that are paid by the investor every time shares in a basket are sold, ETFs avoid these costs. This means that investing via index funds is associated with increased explicit (e.g., commissions) and implicit (trading fees) costs resulting from redemptions. In contrast, when an ETF requires redemption, it does not pay in cash for the exited securities but rather with in-kind positions in other securities (what is known as the redemption mechanism) (SEC, 2012: 2). Additionally, to pay for the daily net redemptions, index funds incur costs related to cash drag which increases the overall costs for an investor as well.

Another crucial feature that distinguishes ETFs from traditional open-ended mutual funds is trading flexibility. While open-ended mutual funds can only be bought and sold at the end of the trading day when the NAV is officially announced, trading in ETFs takes place throughout the day.

The valuation of ETF shares is continuous during normal exchange hours. As the pricing of the underlying securities changes as a response to the growing or falling demand, the price of ETF shares is adjusted immediately. Therefore, ETF constitutes an instrument that is appropriate both for long-term and short-term investors.

The trading flexibility of ETFs offers its clients a variety of investment opportunities. Investment in ETF shares allows one to benefit from all the trade combinations typical for common stocks, including limit orders, stop-limit orders, borrowing on margin, and short selling (Ferri, 2009: 60).

3. 3. 3 Transparency

One crucial feature of ETFs is a high level of transparency. This is determined by the rules of their operation, the methods of trading the units, and how they are valued. The operation of

ETFs, like other open-ended investment funds, is strictly regulated by national or regional legislation (e.g. Investment Company Act of 1940 in the US, and the UCITS Directive in the European Union). As the regulations require ETFs to make a wide range of information public and update it regularly, they operate with great transparency. The key information provided by ETFs through their websites and periodic reports is presented in Figure 12.

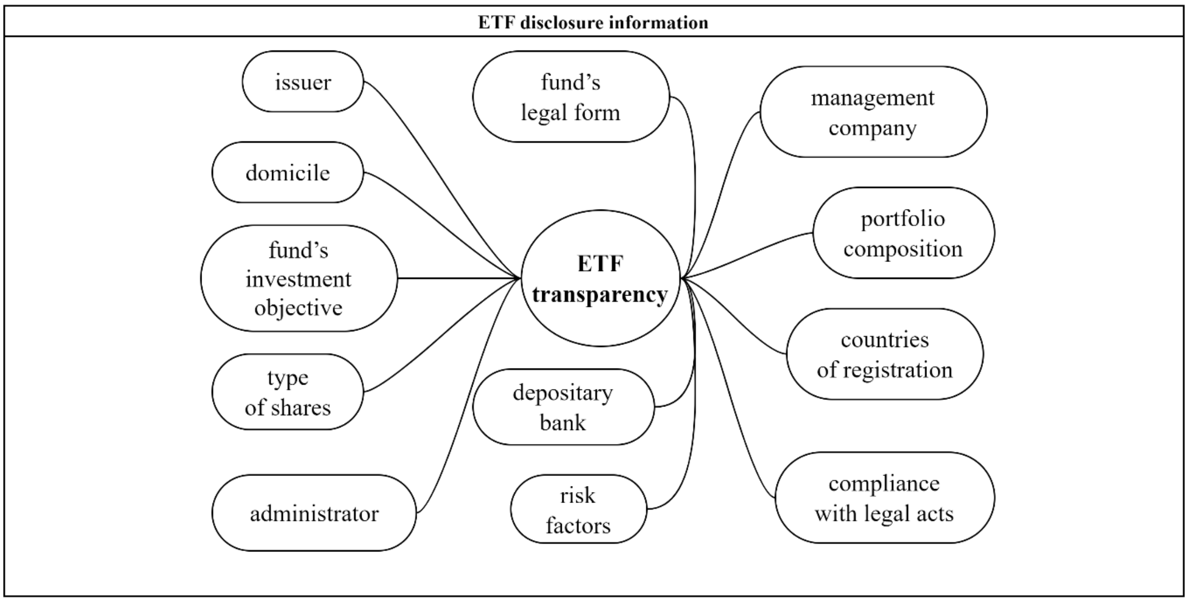

Figure 12. Primary information disclosed by exchange-traded funds

(Source: Own elaboration based on Miziołek, 2013:450)

Moreover, ETFs often publish on their websites a great deal more information relating to (Miziołek, 2013: 451–455):

— quotation of units on exchanges (e.g., primary exchange, stock exchange ticker, initial listing date, market makers);

— trading of units (e.g., trading hours, currencies, trading unit);

— pricing of units (e.g., opening price, previous close, last price size, ask price, ask size, bid price, bid size, spread, day's move);

— book value of shares (e.g., valuation currency, NAV, total return NAV, intraday NAV);

— ETF performance (e.g., total return,

— ETF replication efficiency (e.g., tracking error);

— benchmark index (name, identification data, index provider, index inception date, index construction and calculation methods, index constituents);

— ETF structure (index replication methods, counterparty data, and swap transaction information for ETFs applying synthetic replication);

— securities lending (only for ETFs using physical replication methods)

— fund's net asset value (e.g., outstanding shares, shares in issue, total net assets, assets under management);

— income treatment (accumulative ETFs that reinvest income for investments or distributing ETFs that distribute income periodically through dividends to fund participants)

— taxation of the fund's income (e.g., fiscal year, fiscal year-end);

— participation costs (total expense ratio – TER, management costs).

Particular transparency distinguishes passively managed ETFs that apply the full physical replication method (they constitute a substantial share of the exchange-traded funds' industry). They achieve their primary objective, which is to replicate the performance of the tracked benchmark as closely as possible by investing money collected from clients in the same financial instruments and the same proportions as the composition of the selected index. Consequently, they allocate capital in a basket of securities that is easy to find for clients, as the composition and weights of the components of financial market indices are publicly available and updated on a daily basis. This, in turn, makes them outstandingly transparent.

Synthetic ETFs, by their very nature, exhibit a higher degree of complexity and therefore appear to be less transparent than funds using physical replication methods. As they often invest in a completely different basket of assets than the underlying benchmark, their transparency depends on the extent and frequency of information disclosure regarding the composition and structure of the portfolio. Besides, a common practice among such funds is to publish information on the type of collateral used in the swap agreement, counterparty risk management methods, and the costs involved. On the whole, the more information ETF discloses, the more transparent and trustworthy it is, which results in increased investor interest in purchasing shares of such ETFs.

From the perspective of the principal subject of the paper, which depicts ESG ETFs, ESG disclosure information is essential. ESG-oriented ETFs publish some details about the adopted investment strategy, which determines the investment objective of the fund. This enables ESG investors to evaluate whether the strategy is consistent with their needs. The ESG fund factsheet also includes information about the type of ESG exposure, specific methodology used to select companies, inclusion and exclusion criteria, as well as criteria applied to establish portfolio weightings. Moreover, the information about ESG data and ratings providers matters as a guarantee for the ESG information quality. Then, ESG ETFs commonly publish details related to ESG-related risk factors, including regulatory changes, reputational risks, or the potential for lower returns due to ESG screening. Finally, funds tracking ESG indices constructed on the basis of their parental indices typically refer to the influences of the parental index holdings on the composition of the ESG portfolio, showing both similarities and differences.

Even though ESG exchange-traded funds deliver much more information on ESG-related issues than their non-ESG counterparts, they are still criticised as insufficiently transparent. In fact, it applies not only specifically to ESG funds but also to the ESG industry as a whole. One survey conducted by the Index Industry Association (2022) showed that 64% of asset managers from France, Germany, the United Kingdom, and the United States were concerned about insufficient transparency and corporate disclosure relating to companies' ESG activities. The major concerns are derived from deficient ESG data and measurement, and significant regulatory disconnect.

To improve ESG transparency U.S. Securities and Exchange Commission (2022) obliged ESG funds to define strictly its ESG strategy strictly to differentiate between ESG-focused funds, ESG integration funds, and impact funds. In addition, the Commission indicated that certain environmentally focused funds need to disclose the greenhouse gas (GHG) emissions associated with their portfolio investments. In Europe, ESMA (2024) published guidelines on the use of ESG- and sustainability-related terms in fund names. According to the rules, funds using such terms must allocate at least 80% of their assets to investments promoting environmental or social characteristics or sustainable objectives and must apply exclusion criteria to avoid greenwashing (ESMA, 2024). The introduction of clear and common regulations regarding ESG funds naming and disclosure information would inevitably bring more understanding of ESG ETFs. Importantly, to constitute a real alternative for traditional non-ESG ETFs, the funds need to be as transparent and easy to compare as the traditional ones. However, due to numerous ambiguities in ESG investing signalled in the previous chapter, the issue of ESG ETFs' transparency also requires special efforts since the typical ETF regulations are not sufficient for ESG ETFs.

3. 3. 4 Versatility

Lastly, ETFs benefit investors 'interests due to their universality. The specificity of ETFs made them extremely flexible by serving as a package for multiple securities. Exchange-traded funds offer a wide range of investment opportunities. With ETFs, investors can gain exposure to various asset classes, specific industries, countries, or defined investment styles. This enables the utilization of ETFs for various purposes and investment strategies.

While originally ETFs offered exposure to the broad equity market, with the industry's expansion, more and more sophisticated investment products are available. Importantly, the development of ETFs' product range was a gradual process. Firstly, since 1990, the ETF industry has been dominated by capitalization-weighted broad market indices. The consecutive innovations included the expansion from ETFs tracking domestic indices, to abroad market indices, international equity indices, as well as fixed-income, precious metals, commodity, currency, and real estate indices. Then, since the introduction of smart-beta indices in 2003, the ETF industry started a slow turn into more active portfolio management. The first actively-managed ETF – the Current Yield ETF was launched in 2008 by Bear Stearns, which marked the beginning of the third phase of ETF industry development (StateStreet, 2023). This has made ETFs largely universal products appropriate for both passive and active investors and extended the potential customer base.

Table 9. ETF Investment Universe

	INDEX	**NON–INDEX**
ACTIVE	Alpha-Seeking Enhanced Indices	Fundamental Model-Driven Active
PASSIVE	Capitalization-Weighted Index	Non-Index Beta Exposure

(Source: Golub et al., 2013: 6)

As shown in Table 9, ETFs can be considered from two dimensions: active vs. passive and index vs. non-index. Firstly, traditional passive index ETFs simply track the performance of a capitalisation-weighted index. Then, active index-based ETFs intend to outperform the performance of a capitalisation-weighted index by applying special algorithms designed to generate alpha or enhanced risk-return characteristics. The rules include overweighting assets with low P/B value, dividend stocks, growth stocks, etc. Next, active non-index-based ETFs in many ways resemble active mutual funds. They aim to outperform the broad market by actively selecting assets for a portfolio overseen by a professional asset manager. Their unquestionable advantage is the flexibility to react quickly to market changes and benefit from sudden events, while their disadvantage entails high costs. Finally, some ETFs follow a passive strategy, so they aim to mirror the performance of an asset as closely as possible, but do not track any index. For example, exchange-traded products providing exposure to a currency or precious metal often physically hold the benchmark asset, so there is no need to track an index even though it is a passive investment (Golub et al., 2013: 5–6).

With such a wide range of investment opportunities offered by the ETF industry, the product is extremely flexible in terms of its applicability for both strategic portfolio management and tactical changes (HKEX, 2017: 23-24). Fundamentally, ETFs serve as a long-term asset allocation vehicle. Not only does it allow for desired exposure to particular asset classes, such as equities, bonds, or commodities, but also certain tilts related to company capitalisation or geographical exposure. In addition, ETFs enable portfolio completion by acquiring more sophisticated funds that follow a specific investment strategy or give exposure to a particular market segment. This comprises smart-beta ETFs, sector funds, thematic funds, including ESG ETFs. Thereby, ETFs allow the creation of a unique investment portfolio that is compatible with the investor's individual needs. Finally, the specificity of ETFs enables liquidity management. As ETF units can be sold at any time on an exchange, they can be applied to supply the investor with cash in case of emergencies.

As far as tactical applications of ETFs are concerned, they allow for fast changes and refinements in portfolio composition as a reaction to changes in market conditions. This includes both mitigating risk in response to crises and benefiting from downturns through the use of inverse ETFs. In addition, leveraged ETFs offer the chance to multiply returns (mainly in short-term investments). Then, during the periods of portfolio transition, ETFs might be used to ensure that the portfolio does not miss market opportunities. Finally, ETFs work well as a tool for cash equitization. During transition periods, they facilitate the investment of surplus cash to maximise returns or minimise cash drag.

3. 4 Tracking Performance of ETFs

Recognising that the vast majority of exchange-traded funds are index or passive products, it is time to discuss the principal issue relating to ETFs' tracking ability. As the objective of any passively managed fund is to replicate the return and risk of the underlying benchmark as closely as possible, high tracking ability is the primary investment purpose of such funds.

As simple as index-based products might seem to be, they are not. Frino and Gallagher (2001, 2002) argued that an index represents just a paper portfolio. Passive exchange-traded funds and any other index-based products are created to make the paper portfolio a real investment product. Bearing costs related to the incorporation of passive fund management and market frictions, the flawless mimicking of the benchmark index is unattainable. Most passive ETFs do not fully replicate their underlying benchmarks since their portfolio holdings are not perfectly the same as the index components. Instead, the basket of securities of an ETF is designed in a way that ensures close correlation with the return characteristics of the benchmark.

There are several reasons why ETFs do not hold all the index components in their portfolios. Firstly, it is inefficient to purchase minor illiquid securities constituting a small part of the benchmark portfolio. Secondly, regulatory constraints regarding the maximum weights of single securities prevent ETF managers from fully replicating an index. For instance, the US tax diversification rule obliges ETFs not to hold more than 25% of assets in the largest portfolio holding and requires that the sum of all the holdings constituting more than 5% of assets not exceed 50%. The European UCITS rule, in turn, prohibits an ETF from holding more than 5% of its assets in securities issued by a single issuer. However, it allows this limit to be increased, provided that the total exposure to these issues does not exceed 40%. Finally, a fund manager may prefer to achieve other investment goals, such as enhanced liquidity of an ETF at the expense of lower tracking ability (Golub et al., 2013: 21–22). The following section investigates the key issues relating to the measurement of an ETF tracking accuracy, factors determining the level of tracking ability, and academic research on the topic.

3. 4. 1 Metrics for Evaluating Tracking Accuracy

The measurement of passive ETFs' tracking ability is essential for the assessment of their efficiency for several reasons. First and foremost, this enables us to evaluate ETFs' ability to achieve their primary investment strategy goal. Investors purchasing passive ETF shares wish to obtain a specific market or segment exposure. Exchange-traded funds whose returns significantly diverge from the returns of the underlying index fail to achieve this goal. Consequently, this leads to difficulties in implementing a passive investment strategy. Furthermore, measuring tracking ability provides a standardized measure to compare different ETFs that track the same or similar benchmarks. Among multiple ETF product options, investors might go for the most efficient, which means that the underperformance in relation to the benchmark is the lowest. The basic measures of a passive fund's tracking ability include tracking error and tracking difference, whose interpretation is quite straightforward. The lower the tracking error (and tracking difference), the higher the tracking ability. Below, I will describe the most popular methods of measuring the effectiveness of passively managed funds, with their pros and cons.

It was Roll (1992) who stated that the level of tracking error can be an important criterion for evaluating the performance of an index fund, as the differential return of a fund can indicate whether the investment process has been implemented successfully. The tracking error (TE) depicts the volatility of differential returns between the fund and its corresponding benchmark. In comparison to the tracking difference, it does not provide information about the absolute difference in performance between an ETF and its benchmark. It focuses on the volatility of differential returns. Thus, tracking difference is of great importance in the long term, while tracking error, which informs about the variability of tracking difference, is crucial in short periods. Depending on the purpose of the tracking measure calculation, one might apply historical or predictive data.

According to Pope and Yadav (1994), tracking error can be defined in three different ways: the average absolute difference between an ETF's returns and its benchmark returns, the standard deviation of return differences, and the standard deviation of the residuals in the estimate of the Capital Asset Pricing Model (CAPM). This is the approach selected for conducting the empirical study within this dissertation, and thus, this is thoroughly discussed in section 4.2.

Another measure that can be used to assess an ETF's ability to track the performance of an index is the correlation coefficient of the returns of the ETF and the returns of its underlying benchmark index. This measure quantifies the statistical relationship between the variables and takes values from -1 to 1. The direction of this relationship is determined by the sign of the correlation coefficient (positive or negative), while the strength is expressed by the absolute value of the coefficient (Boddy and Smith, 2009: 95–96).

Bearing in mind that a passive ETF seeks to best reflect the returns of the benchmark index, the highest possible positive correlation (ideally equal to 1) is desirable. This means that an increase (decrease) in the return of the index by a unit is accompanied by an increase (decrease) in the return of the ETF of the same strength and direction. Negative values of the correlation coefficient of index and ETF returns would mean that the manager is acting contrary to its primary objective, i.e. that, for example, an increase in index returns would be accompanied by a decrease in ETF returns. Such situations are unlikely to occur with passive funds. As simple and standardized as this measure might seem to be, the primary limitation of this approach to the tracking error calculation is the lack of information on the magnitude or the magnitude of deviations between returns of the ETF and its benchmark.

The last measure that can be used to assess the effectiveness of passive ETF management is the coefficient of determination, also known as the R^2 coefficient (R-squared). It provides information on how much of the variability in the explanatory variable (ETF return) was explained by the explanatory variable (index return) in the linear model. This measure takes values from 0 to 1, with 1 indicating a perfect fit of the regression line to the empirical data (Devore, 2011: 508–510). It is calculated as the square of the correlation coefficient.

Even though the measures of ETFs' tracking ability might seem to be easy, accurate calculation is connected to some challenges. One challenge derives from data errors, as due to the magnitude of data, ETF issuers may commit mistakes associated with reporting the fund’s NAV. What is more, the errors might occur as a result of holidays, missing or misaligned data, outliers, or rounding. Finally, it is essential to make sure that the ETF’s NAV is juxtaposed against an appropriate benchmark index. There are many different variations of indices depending on methodology or the base currency. For instance, if the fund whose NAV is calculated on a total return basis is compared to a price return index, the divergences would be the implication of choosing the wrong index, not the weak tracking ability of an ETF (Johnson et al., 2013: 8).

3. 4. 2 Research on ETFs Tracking Ability

Research on the tracking ability of ETFs has been conducted since the early 2000s, with the first such study focusing on Spider representing participation in the SPDR Trust (Elton et al., 2002). The authors examined the annual performance of Spider relative to the S&P 500 Index, proving that the NAV and fair market value were very close. However, an ETF underperformed the market by 28 basis points per year. Then, they indicated that the management fee and costs related to non-accruing earnings on dividends were the principal reasons for the underperformance. Similarly, Gallagher and Segara (2005) explored six ETFs listed on the Australian Stock Exchange and showed that all the funds delivered investors a return consistent with the underlying benchmark before costs. In addition, they noted that the variation between ETFs’ NAV and trading price was insignificant and infrequent.

Further evidence of considerable tracking error of ETFs was provided by Milonas and Rompotis (2006), who explored 36 Swiss ETFs. They reported that the average tracking error of the funds under study was 1.02% and statistically significant. They confirmed that an increase in tracking error is influenced by management and risk fees of ETFs, with a one-unit increase in costs reducing the return of ETF by 0.35 units.

Then, Shin and Soydemir (2010) investigated 26 U.S. ETFs, showing that their tracking errors were significantly different from zero, which suggests that the ETFs do not mimic the corresponding indices perfectly. A study by Rompotis (2011) covered a sample of 50 passive ETFs belonging to the family of Barclays iShares over the period 2002–2007. The Author estimated an aggregate tracking error using both ETFs' trading prices and their NAVs. Then, using a regression analysis, he proved that tracking errors are strongly persistent in the short term.

The literature on ETFs' replication quality is quite rich, and describing all the individual research misses the point. Conversely, gauging the results of meta-studies would be more conclusive. Thus, Gaba and Kumar (2021) investigated 57 articles and summarized that the majority of authors pointed to the underperformance of ETFs in comparison to their benchmarks.

Another branch of research concerning ETF tracking ability focuses on different kinds of comparisons. A study by Wong and Shum (2010) showed how the tracking ability of ETFs changes in different market conditions. The authors explored 15 ETFs in the USA, UK, Hong Kong, Japan, Netherlands, and Belgium and concluded that no matter whether it's a bullish or bearish time, ETFs display positive tracking errors.

Some studies compared the tracking efficiency of ETFs with comparison of index mutual funds. According to Rompotis (2005), who investigated the performance of 16 ETFs and index funds in pairs tracking the same index, most often they produce similar returns and tracking errors. Harper et al. (2006) focused on the comparison of the risk and return performance of ETFs to closed-end funds, as well as on the tracking error verification of ETFs. They utilized a sample comprising 29 closed-end funds for 14 countries over the period from April 1996 to December 2001. Their study proved the good tracking ability of ETFs as compared to the benchmark. Then, they delivered some evidence that ETFs exhibit higher average returns than country closed-end funds and pointed to lower expense ratios as the reason for this. A few years later, Rompotis (2009) repeated a study comparing the tracking performance of ETFs and index funds on a sample of 20 Vanguard index funds and 12 Vanguard ETFs. This time, he showed a slight tracking outperformance of ETFs. While the average tracking error of ETFs was 12 basis points, for index funds it was 14 basis points.

A study by Svetina and Wahal (2008) showed that the tracking error of ETFs tracking non-domestic indices was up to double that of domestic counterparts. According to Chu (2011), Hong Kong ETFs exhibited significantly poorer tracking performance than U.S. and Australian ETFs. Similarly, Johnson (2009) and Blitz and Huij (2012) researched ETFs on emerging markets and showed that their tracking errors are much higher than those of ETFs tracking developed markets' indices.

All in all, based on the previous studies, it might be said that discrepancies in return and risk characteristics between ETF and its underlying benchmark are unavoidable. Subsequent authors reported greater or lesser levels of tracking error and tracking difference among ETFs. Nevertheless, none of the explored research pointed tremendously poor tracking ability of ETFs. This demonstrates that, despite some limitations, in general, passive ETFs fulfil their primary purpose quite well.

3. 4. 3 Factors Influencing Tracking Performance of ETFs

As shown above, the complete elimination of tracking error in passively managed ETFs is not possible, and the return of a passive ETF is not likely to be equal to the return of an index. This is because an index is merely a calculation derived from a portfolio of securities that is not affected by the same factors as an ETF. There are several commonly recognized determinants of an ETF's tracking ability that will be explained below.

Firstly, it should be noted that not every factor determining ETF tracking ability is its direct source. Some factors explain the replication quality even though they do not harm or improve it. Table 10 presents the direct causes of tracking differences according to Johnson et al. (2013: 9) in the division of effects on physical and synthetic ETFs.

Table 10. Direct causes of the ETF tracking difference

Cause	Physical ETFs	Synthetic ETFs
Total Expense Ratio (TER)	Negative	Negative
Transaction and rebalancing costs	Negative	N/A*
Cash Drag	Negative / Positive	N/A *
Taxes on dividends and earnings	Negative / Positive	N/A *
Reinvestment delays	Negative / Positive	Negative / Positive
Income from securities lending	Positive	N/A *
Swap Spread	—	Negative / Positive
Rebalancing interval	Negative / Positive	—
Optimised replication	Negative / Positive	—

* Note: It considers the swap spread
Source: Johnson et al. (2013: 9)

One widely known factor determining tracking ability is the fund's expense ratio. The positive relationship between the fund's expense ratio and tracking error was documented by numerous authors, including Frino and Gallagher (2001), Rompotis (2009), Agapova (2011), Blitz et al. (2012), Elia (2012), and Osterhoff and Kaserer (2016). All of them indicated that the higher the expense ratio, the lower the tracking ability. Charupat and Miu (2013) showed that with an increase in an ETF's expense ratio, the expected underperformance to the underlying index raises.

The most common measure of ETF costs in the literature is the total expense ratio (TER), which represents the annual cost of owning an ETF. It constitutes the explicit costs of managing an ETF and represents the expense scaled by the size of an ETF. TER of an ETF is expressed as a percentage of the fund's assets under management (AUM) and is deducted from the fund's net asset value (NAV) on an ongoing basis. This means that an ETF's performance already includes the impact of fees, so when the total expense ratio (TER) increases, it becomes harder for the ETF to match the return of its benchmark index. Essentially, the total expense ratio encompasses the management fee, operating costs, and other costs associated with running an ETF. However, it should be emphasized that these are not all the costs incurred by an investor purchasing ETF units. For instance, the total expense ratio does not include transaction costs and taxes, which also affect the final rate of return that an investor achieves from holding ETF units.

Next, it is important to consider what drives the differences in the total expense ratio across ETFs. Primarily, management costs are mostly responsible for the level of TER. Management costs are reflected in the valuation of ETF units as they are deducted from the fund's assets. This has a detrimental effect on the fund's performance and prevents the fund from investing in the index the full amount of capital collected from its clients (the phenomenon known as the cash drag). Moreover, depending on the asset class an ETF invests in, the costs might be higher or lower. For example, ETFs tracking typical asset classes and developed markets tend to have lower TER. Investing in emerging markets assets or less liquid securities involves additional expenses such as higher trading costs or bid-ask spreads.

Then, one common conviction is that ETFs using the synthetic replication method have lower tracking errors than ETFs applying physical replication. Elia (2012) investigated 48 ETFs from the leading providers that track major European and global stock market indices over the period 2007–2011 and proved that ETFs using synthetic replication were rewarded with lowered tracking errors, especially those tracking emerging market indices. Similarly, Johnson et al. (2013), who researched 65 ETFs from European countries confirmed that ETFs with physical replication had higher tracking errors than ETFs applying synthetic replication. This is because synthetic ETFs do not incur rebalancing costs. Instead, the ongoing charges and swap spreads constitute a significant share of their costs, which most often are lower than costs related to rebalancing.

Then, Buetow and Henderson (2012) studied around 3,700 ETFs traded on U.S. markets across all asset classes to show that the liquidity of the underlying assets strongly affects ETFs' tracking error. The trading volume of ETFs was used as a proxy for their liquidity. High-yield corporate bonds proved to be the most illiquid assets, which was mirrored in the poorest replication quality of ETFs tracking indices of such securities. Furthermore, during times of crisis, the tracking ability of funds investing in illiquid assets was particularly strongly affected.

Another factor influencing ETFs' costs and, consequently, their tracking error is ETF size – the amount of total assets of a fund. Chu (2011) pointed out that size is negatively related to tracking error since larger ETFs may face lower transaction costs resulting from economies of scale. Transaction costs encompass the explicit costs of trading in stock markets, like brokerage fees and stamp duties. While indices are computed based on the assumption that transactions are free of costs, ETFs are required to trade in the markets, which implies incurring costs. The costs can significantly reduce ETFs' returns, and consequently, the ability to replicate the index performance.

Then, Chu (2011) evidenced that tracking errors in Hong Kong ETFs are negatively related to size but positively related to dividend yield, trading volume, and market risk. Larger ETFs can more easily replicate the index because they are more liquid, have better access to the markets, and may benefit from securities lending. Then, due to their limited resources, smaller ETFs might struggle to keep an appropriate mix of index constituents.

Then, Frino and Gallagher (2001) identified dividend payment as a possible factor contributing to a decrease in ETFs' tracking ability. While indices assume immediate reinvestment of dividends in the stocks on the ex-dividend day, fund managers receive dividends with some delay. Additionally, funds incur transactional costs, which lead to returns depletion. The two factors might considerably decrease ETFs' ability to replicate an index. Elton et al. (2006) pointed out the delay in reinvestment of cash dividends as one of the major causes of tracking error. Blitz et al. (2012) and Blitz and Huij (2012) argued that dividend taxes explain ETFs' expected returns and their tracking error. Finally, De Fusco et al. (2011) said that the accumulation of dividends by ETFs impacts tracking errors measured by pricing deviation. Accordingly, research shows that the higher the dividend yield, the time delay in receiving it, and the return of the underlying index, the lower the ETF's tracking ability.

Moreover, according to De Fusco et al. (2011), the number of constituent stocks in an index might determine an ETF’s tracking ability. Because the exact proportions of the constituent stocks in the index are not public, fund managers have to work out the portfolio structure on their own. Thus, a growth in the number of constituent stocks in the index increases the difficulty of mimicking the index. This implies a positive relationship between the number of stocks in the index and the ETF’s tracking error.

Shin and Soydemir (2010) applied a panel regression model approach for U.S., American, Asian and European ETFs and explored the significance of five factors: expense ratio, volatility of ETFs’ daily market prices, the natural logarithm of average daily trading volume, returns on the dividend from the stocks each ETF owns in the portfolio and exchange rate. Interestingly, only the exchange rate proved to be a statistically significant determinant of ETFs' tracking error. The authors explained that an increase in the exchange rate corresponds to a depreciation of the U.S. dollar, so the conversion from local currency to U.S. dollars leads to an increase in the net asset value of ETFs. This, in turn, increases the dispersion between the ETF NAVs and benchmark returns, causing higher tracking errors. This is why the U.S. ETFs benefited from the lowest level of tracking errors while American, Asian, and European ETFs exhibited higher levels of tracking errors.

Rompotis (2011) showed that expenses, risk, and age are factors that can explain the persistence in tracking error for exchange-traded funds. The positive relationship between market volatility and tracking error was also evidenced by Qadan and Yagil (2012) and Drenovak et al. (2014). The phenomenon is related to the growing difficulty of replicating the market performance as the market is more volatile, which leads to a higher tracking error. Thus, with growing market volatility, ETFs’ tracking error rises, and, conversely, when the market is more stable, the tracking error is smaller.

Some studies show that tracking error in ETFs often follows an autoregressive pattern. This reflects that when fund managers find effective ways to track an index, they continue to use them. DeFusco et al. (2011) found that these patterns can be described with special models like ARCH or GARCH. Ivanov (2015) showed that such models are appropriate to model the tracking errors of currency ETFs.

3. 5 Summary

This chapter has outlined the fundamentals of ETF operation, with a particular focus on their tracking ability. ETFs have become extremely popular with investors around the world for several reasons. In addition to the advantages typical of all funds, like providing diversification, they are characterized by high liquidity, transparency, and, above all, low costs.

What is unique to ETFs is that they operate both on the primary and secondary markets. ETF units can be traded on the market like shares without worrying about liquidity. Central to ensuring their liquidity is the role of Authorized Participants, who create or redeem ETF units accordingly in the event of unbalanced supply and demand. The creation and redemption mechanism enables ETFs to offer clients lower fees than index mutual funds.

Most ETFs are passively managed, so their principal aim is to track the return of the underlying indices. ETFs' tracking ability remains the key concern in their management. Investors care about the replication quality of ETFs as the deviations lead to breaches of passive investing premises and diminished returns. The studies on ETFs' tracking ability demonstrate that exchange-traded funds typically underperform their benchmarks, however, the divergencies are not usually major. The reviewed research indicates that the tracking ability of ETFs varies depending on the underlying asset class, geographical region, and market volatility. The key determinants of tracking error include transaction and rebalancing costs, cash drag, the tax treatment of dividends, fund size, and the replication method. Market volatility, exchange rate fluctuations, and ETF age also play a role. Finally, tracking error often exhibits autoregressive properties, suggesting persistence and predictability over time.

CHAPTER 4 TRACKING ABILITY OF PASSIVE ESG EQUITY ETFs – EVIDENCE FROM EUROPEAN MARKET

4. 1 Introduction

The preceding chapters reviewed the literature on passive portfolio management, ESG investing, and ETF tracking ability. This chapter presents the empirical analysis designed to address the dissertation's core objective: to evaluate and compare the tracking performance of passive ESG equity ETFs listed on European exchanges with their non-ESG counterparts, and to identify the key determinants influencing replication quality.

The study adopts several methods, including descriptive statistics, significance tests, correlation analysis, regression, and dynamic panel data techniques. The dataset includes 134 passively managed ETFs, and the period of analysis spans from January 1, 2021, to June 30, 2024.

This study uses comparative analysis to verify differences in tracking errors between ESG and non-ESG ETFs. Panel linear regression models are applied to explore how the past tracking error, total expense ratio, assets under management, age, the replication method, and the ESG variable impact tracking error. The regression analysis is conducted for the full sample of ETFs, and then for ESG and non-ESG ETFs separately. This approach aims to reveal both general patterns across all ETFs and unique influences within the fund categories. The chapter is organized as follows. The first section provides an overview of the research's methodology and scope. The next section deals with sample characteristics. Ultimately, the chapter summarizes and evaluates the findings, concluding with the impact of ESG integration on the tracking ability of passive equity ETFs.

4. 2 Research Design and Methods

The following section explains the procedures used to validate the research hypotheses. The research uses quantitative methods, ensuring that the findings are comparable with existing studies. (Zheng, 2021; Dyer and Guest, 2022). By using a top-down approach, the research is grounded in existing theories on the tracking ability of ETFs. It develops the knowledge of ETF tracking ability by transferring well-established dependencies from non-ESG ETFs to ESG ETFs.

Key methodological challenges stemmed from the selection of the research sample and study period. While a global analysis could offer a broader perspective, cross-market differences in regulation, liquidity, and currency exposure would hinder comparability. The study focuses on the European ETF market as a single, consistent market. There are several reasons for this choice. First, Europe is a leader in the global ESG ETF market. At the end of 2023, Europe managed nearly 75 percent of global ESG ETF assets, totaling 402 billion dollars (Trackinsight, 2024). Second, the European regulatory framework provides exceptional support to ESG investment activities. The Sustainable Finance Disclosure Regulation (SFDR), together with the EU Taxonomy Regulation and other regulations discussed in Chapter 2, have established specific standards for financial institutions. Finally, the regulatory clarity enhances transparency, which enables better results comparison.

The comparative part of the research required creating matched groups of ESG and non-ESG ETFs. Without making the groups similar, any differences in results could be caused by other factors, not by the impact of the ESG factor. To address the problem, the study incorporates numerous inclusion and exclusion criteria that make the sample homogenous. Funds that use leverage, inverse strategies, or currency hedging were excluded from the study because these approaches create different levels of risk, returns, and costs, which can strongly influence tracking performance. Also, since exchange rate changes affect tracking error, only funds using the same currency as their benchmark index were included.

A key rule for including ESG ETFs in the sample was the existence of a non-ESG counterpart tracking an index with the same regional exposure. This assures that the differences in tracking performance are due to the ESG factor, not differences in the regions the funds invest in. Consequently, ESG and non-ESG counterparts in the sample track indices with the same geographical exposure delivered by the same provider, making the ESG factor the primary differentiator between funds. Minor discrepancies in index return types (e.g., gross return, net return, price return) were accepted to preserve an adequate sample size.

The analysis of ESG ETFs introduced certain constraints because their historical data was limited. ESG ETFs represent a newer investment product than traditional equity ETFs, as they entered the market more recently. Before 2021, there were not enough ESG funds available for analysis. To ensure the data is complete, the study includes only funds that started before January 1, 2021, and were still active on June 30, 2024. This approach includes many ESG ETFs launched during the early months of the COVID-19 pandemic and provides a solid research period of 42 months. Table 11 summarizes the adopted criteria for the sample selection.

Table 11. Summary of inclusion and exclusion criteria in the research sample

Inclusion criteria	Exclusion criteria
fund listed in the European market	fund established after January 1, 2021
passively managed fund	leveraged fund
broad-market equity fund	inverse fund
fund currency consistent with the currency of the benchmark index	fund adopting currency hedging
existence of an ESG/non-ESG counterpart	

Source: Own elaboration

The side effect of the adopted approach is a considerable reduction in the sample size. As shown in Figure 13, the number of ETFs under study represents approximately 10% of the basic sample. However, given the comparative nature of the study, this procedure was necessary to obtain reliable results.

The research covers the period from January 1, 2021, to June 30, 2024, and is supported by the following arguments. Firstly, starting the research on January 1, 2021, has significantly increased the likelihood that the ESG ETFs under study apply a transparently defined ESG strategy. It was on March 10, 2021, when the Sustainable Finance Disclosure Regulation (SFDR), representing a critical part of the sustainable finance regulations in Europe, came into force. Before the SFDR's implementation in 2021, many ESG ETFs in Europe were criticized for engaging in greenwashing. This means that ETFs were marketed as ESG-compliant without significantly differing from non-ESG ETFs, except in name. Following Cohran et al. (2024), this study assumes that even though this regulation has not fully eliminated the greenwashing problem, it has contributed to considerable progress in mitigating greenwashing.

Secondly, as shown in Figure 14, the number of ESG ETFs and their assets grew significantly after the COVID-19 period. Before this period, the number of such funds was limited, which would have substantially constrained the research sample. Starting the analysis in 2021 allows for capturing a sufficiently large and representative sample of ESG ETFs.

Finally, the dataset spanning from January 2021 through June 2024 ensures complete data availability, thus reducing the bias and errors that occur when data needs to be estimated. Therefore, this study prioritizes data reliability and completeness rather than artificially extending the research period.

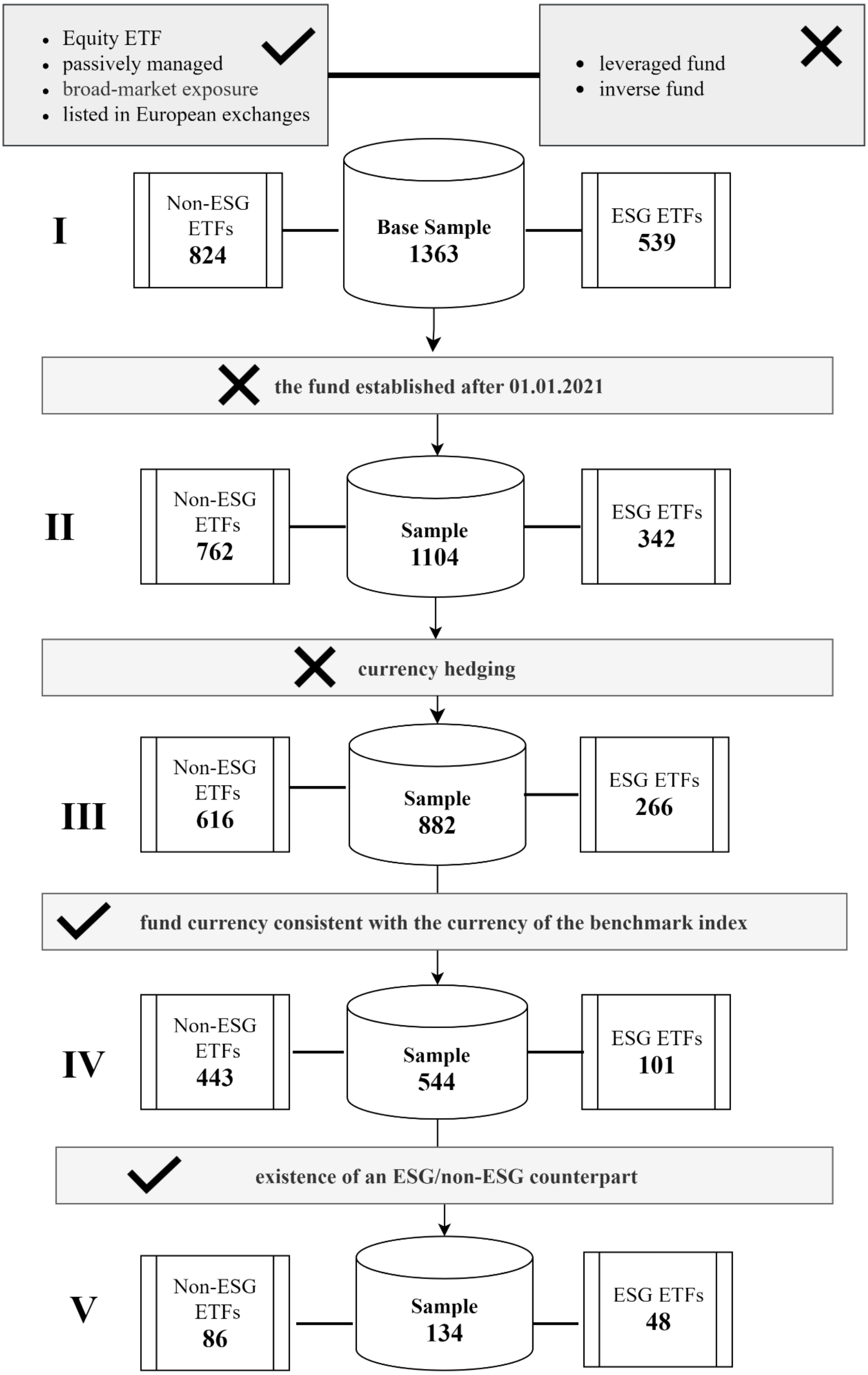

Figure 13. Sample Selection Process

(Source: Own elaboration)

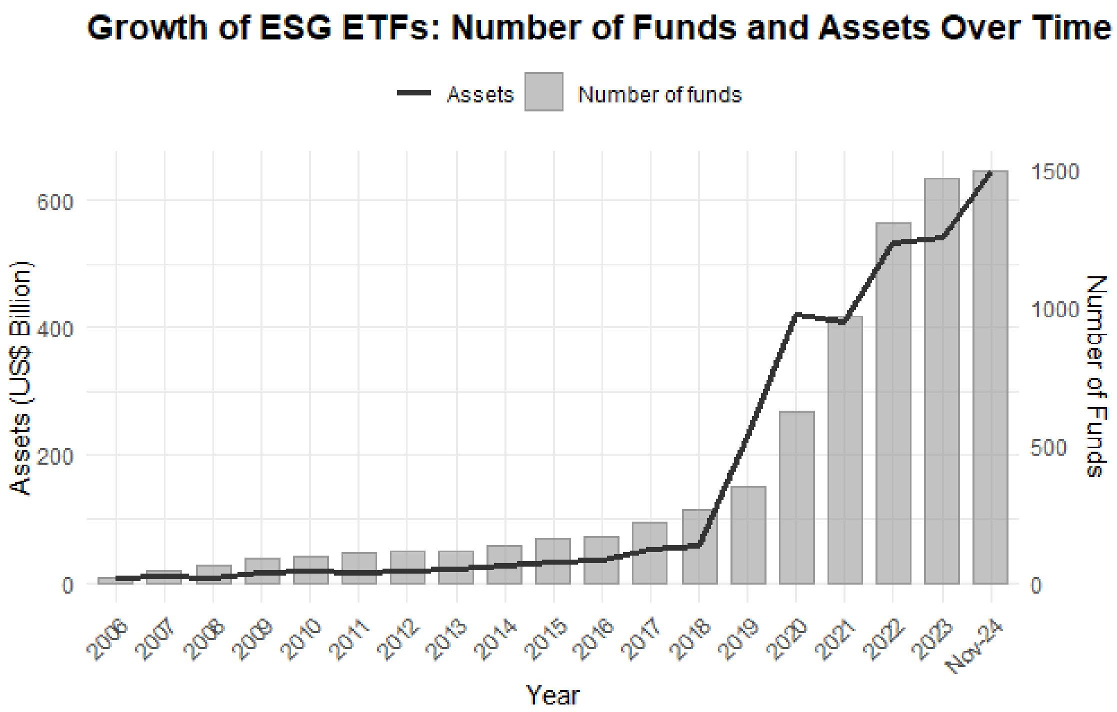

Figure 14. Growth in ESG ETF Assets and Number of Funds (2006-2024)
Source: Own elaboration based on ETGI (2025)

The data for this study was gathered from the Bloomberg Terminal. It provides the clear division of ETFs into ESG and non-ESG categories, and covers a wide variety of financial instruments, guaranteeing the comprehensiveness of the data used in this study. The tracking ability of ETFs is calculated based on weekly data. Therefore, the closing prices of benchmarks and the NAVs (net asset values) of exchange-traded funds have been downloaded from the Bloomberg Terminal on a weekly basis. The study does not employ the trading prices of ETFs, since investors are interested in the actual value of their investments expressed in NAVs. The weekly frequency effectively captures significant trends and deviations without being overwhelmed by short-term fluctuations, making it a well-regarded choice in the literature. For instance, studies by Frino and Gallagher (2001) and Rompotis (2009) indicate that using weekly data balances the need for detailed performance analysis while accommodating the availability of key financial metrics. This frequency provides enough observations to compute reliable tracking error metrics and mitigates the impact of market volatility.

The selection of variables is based on previous studies confirming their impact on the tracking error of non-ESG ETFs (Blitz et al., 2012; Poterba and Shoven, 2002; Elton et al., 2004; Johnson and Gubellini, 2011; Dyer and Guest, 2022). The variables and abbreviations used in the study are explained in Table 12.

Table 12. Explanation of variables used in the study

Variable symbol	Description
TE_1	Dependent variable, tracking error of an ETF calculated according to the first method (the average absolute difference between an ETF's returns and its benchmark returns)
TE_2	Dependent variable, tracking error of an ETF calculated according to the second method (standard deviation of the return differences between an ETF and its benchmark)
TE_3	Dependent variable, tracking error of an ETF calculated according to the third method (standard deviation of the residuals from a regression of the ETF return on the benchmark return)
TER	Independent variable, total expense ratio of an ETF: the annual cost of holding an ETF, expressed as a percentage of the fund's average net assets. This is a measure of the total costs associated with managing and operating an investment fund, TER includes management fees, administrative fees, operating costs, custodian fees, and registration fees.
AUM	Assets under management of an ETF. This is a key metric to assess the size of the ETF and refers to the total market value of the assets held within the ETF's portfolio (securities, cash holdings, derivatives, commodities, and any other financial instruments the ETF might hold).
AGE	Natural logarithm of an ETF age, the difference between the current date and the inception date
RISK	Benchmark risk calculated as the standard deviation of an index monthly logarithmic returns
REP	Independent dummy variable equal to 1 if an ETF uses a synthetic replication method
ESG	Independent dummy variable equal to 1 if an ETF tracks an ESG index

Source: Own elaboration

Each independent variable addresses a different aspect of an ETF's performance. While AUM and fund age offer information about the fund's operational scale and experience, the TER encompasses direct information on the fund's cost efficiency. The replication method clarifies the fund's approach to index tracking, while benchmark volatility captures the inherent market risk. Finally, the ESG variable provides information about the underlying index version.

The successive phases of the research are highlighted in Table 13. The study includes multiple tracking error measures to capture different aspects of ETF replication quality. Given the comparative nature of the study, it verifies the significance of differences using statistical tests and regression analysis. Panel regression techniques are used to identify determinants of tracking error, including ESG-specific effects through interaction terms. Diagnostic testing and robustness checks ensure model validity and result stability.

The first step in measuring ETF tracking error involved the calculation of returns for both the ETFs and their respective benchmark indices. Logarithmic weekly returns were computed using the following formula:

$$R_{i,t} = \ln\left(P_{i,t} + D\right) - \ln\left(P_{i,t-1}\right) \tag{1}$$

where:

$R_{i,t}$ – logarithmic return of an ETF in *t* period,

$P_{i,t}$ – NAV price of an ETF in *t* period,

D – dividend paid during the t period (if applicable),

$P_{i,t-1}$ – NAV price of an ETF in *t –1* period.

Analogously, logarithmic returns of ETF's underlying benchmarks were calculated as follows:

$$R_{b,t} = \ln\left(P_{b,t}\right) - \ln\left(P_{b,t-1}\right) \quad (2)$$

where:

$R_{b,t}$ – logarithmic return of the benchmark index in *t* period,

$P_{b,t}$ – closing price of the benchmark index in *t* period,

$P_{b,t-1}$ – closing price of the benchmark index in *t –1* period.

Following Pope and Yadav (1994), tracking errors were calculated using three methods: absolute return difference (TE_1), standard deviation of return differences (TE_2), and standard deviation of the residuals from a regression of the ETF return on the benchmark return (TE_3). Firstly, tracking error denotes the average absolute difference between an ETF's returns and its benchmark returns. This approach assumes that any deviation from the benchmark return represents a tracking error for the fund. It is calculated as follows:

$$TE_{AD,i} = \sum_{t=1}^{n} \frac{\left|R_{i,t} - R_{b,t}\right|}{n}; \quad (3)$$

where:

$TE_{AD,i}$ – tracking error (as the average absolute difference between an ETF's returns and its benchmark returns),

$R_{i,t}$ – the return of the ETF *i* in the period *t*,

$R_{b,t}$ – the return of the benchmark index *b* in the period *t*.

n – the number of periods.

The tracking difference $\left|R_{i,t} - R_{b,t}\right|$ informs about the difference in the performance of the ETF and the underlying benchmark index over a specific period. It shows how much the fund's return deviates from its benchmark performance and therefore how well the passive ETF is meeting its primary investment objective. This method is easy to use, but it only shows the average return differences and does not account for volatility. It is usually used in conjunction with the other measure for a more complete assessment.

Table 13. Summary of research steps and applied methods

<table>
<tr><th>Stage</th><th>Method</th></tr>
<tr><td>1: Data Collection</td><td>Identification of inclusion and exclusion criteria to compile a list of passive broad market ESG equity ETFs and their non-ESG counterparts listed in European exchanges.
Gathering data on ETFs NAVs, TER, AUM, launching dates, replication method, and benchmark closing prices.
Preparing independent variables: calculating ETFs' age, conversion of ETFs' AUM to USD, and coding binary variables.
Tracking error calculation for both ESG and non-ESG ETFs over a specific period using three methods (TE_1: absolute return difference between the ETF and its benchmark index, TE_2: standard deviation of the return differences, TE_3: standard deviation of the residuals from a regression of the ETF returns on the benchmark return).</td></tr>
<tr><td>2: Descriptive statistics</td><td>Calculating the summary statistics (mean, median, standard deviation, minimum, maximum) for each variable, a preliminary comparison of variables describing ESG and non-ESG ETFs</td></tr>
<tr><td>3: Testing H1: Passive ESG equity ETFs listed on European exchanges exhibit significantly different tracking errors compared to their non-ESG counterparts.</td><td>Conducting the T-test/ Mann-Whitney U Test/ Wilcoxon rank sum test to assess the significance of differences in TE_1, TE_2, and TE_3 between the two groups based on the cross-sectional data, and performing regression analysis including the ESG variable in the panel linear model.</td></tr>
<tr><td>4. Testing H2: The determinants of tracking error differ between passive ESG equity ETFs listed on European exchanges and their non-ESG counterparts.
Testing HS1: Tracking errors of both ESG and non-ESG passive equity ETFs listed on European exchanges exhibit an autoregressive pattern.
Testing HS2: There is a positive relationship between the total expense ratio (TER) and the tracking error of both ESG and non-ESG passive equity ETFs listed on European exchanges.
Testing HS3: There is a negative relationship between assets under management (AUM) and the tracking error of both ESG and non-ESG passive equity ETFs listed on European exchanges.
Testing HS4: There is a negative relationship between fund age and the tracking error of both ESG and non-ESG passive equity ETFs listed on European exchanges.
Testing HS5: There is a positive relationship between benchmark volatility and the tracking error of both ESG and non-ESG passive equity ETFs listed on European exchanges.</td><td>Estimation of baseline panel regression model for the full sample of ETFs
Separate regressions for ESG and non-ESG funds
Chow test to examine structural differences between groups.
Model with interaction terms (ESG × continuous variables) to capture ESG-specific effects.
Diagnostic tests:
– Linearity: RESET test
– Residuals independence: Wooldridge test for autocorrelation in panel data, Breusch-Godfrey test
– Homoscedasticity: Breusch-Pagan test
– Normality: Shapiro-Wilk test
– Endogeneity DWH test
– Multicollinearity VIF test
– Stationarity: Levin-Lin-Chu and Im-Pesaran-Shin tests
Evaluating the regression coefficients and their significance
Robustness checks with the use of different tracking error measures</td></tr>
</table>

Source: Own elaboration

The second approach to measure the tracking error comprises the calculation of the standard deviation of return differences between the fund and the benchmark index. By measuring the volatility of a fund’s return against its benchmark this method provides information about the consistency in the fund's replication. Importantly, this approach is consistent with the definition of the European Securities and Markets Authority’s consultation paper (ESMA, 2012: 43) stating that tracking error is the volatility of the difference of the returns of the fund and of the returns of the index. This method requires the assumption that the return differences are not serially correlated. As daily returns are almost certainly serially correlated, this method is not suitable for daily data. Furthermore, one serious drawback of this approach is it doesn’t capture the actual magnitude of underperformance or outperformance of an ETF and in this regard tracking difference is more useful method. Additionally, if the fund regularly underperforms and outperforms the tracked index by the same magnitude, the tracking error would result in zero which is not compatible with reality. Tracking error as the standard deviation of return differences is calculated as follows:

$$TE_{i,t} = \sqrt{\frac{1}{n-1}\sum_{t=1}^{n}\left(\left(R_{i,t} - R_{b,t}\right) - \frac{1}{n}\sum_{t=1}^{n}\left(R_{i,t} - R_{b,t}\right)\right)^{2}}\ ; \tag{4}$$

where:

$TE_{SD,i}$ – tracking error (as the standard deviation of differences in returns of an ETF and benchmark),

The third method of tracking error calculation applies the residuals from the regression of ETF’s return on the benchmark return. The tracking error constitutes the standard deviation from the residuals e_t in the estimate of Capital Asset Pricing Model (CAPM) as follows:

$$R_{i,t} = \alpha + \beta \times R_{b,t} + e_t; \tag{5}$$

where:

α – alfa coefficient (excess return),

β – beta coefficient (fund systematic risk),

e_t – regression residuals.

This approach considers the peculiarities of the returns of the fund that are not explained by the systematic risk factors of the CAPM. It can provide insight into the manner in which the fund's returns deviate from what would be theoretically expected based on its exposure to the market. The limitations of this approach include the reliance on the assumptions of CAPM and overlooking of the non-market factors.

Data used for the regression analysis was transferred to a 0–1 scale by using the min-max scaling method from Han et al. (2011: 51). With a unified scale, different measurement units of variables do not affect values of the regression coefficients. The transformation was performed as follows:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)}; \tag{6}$$

where:

x' – the scaled value of the variable,

x – the original value of the variable,

$\min(x)$ – minimum value of the variable in the dataset,

$\max(x)$ – maximum value of the variable in the dataset.

Then, in the regression model including interactions with the ESG variable, all continuous variables were first scaled to a 0–1 range using the min-max method. Subsequently, the scaled variables were centered to their mean values. Centering the variables improves the interpretation of the interaction coefficients and reduces potential collinearity problems (Aiken and West, 1991). The process was performed according to the following formula:

$$x'^{c} = x' - \overline{x'}; \tag{7}$$

where:

x'^{c} – the centred and scaled value of the variable,

$\overline{x'}$ – the mean of the scaled variable in the dataset.

One limitation of the study entailed the availability of historical data for some of the variables, especially the total expense ratio (TER). Given the relatively short research period, the study assumes that TER remained stable and uses the values as of May 30, 2023, the date on which the sample was constructed. A similar assumption was made by Rompotis (2011), who noted that ETF expense ratios tend not to vary significantly over short periods.

The panel data regression was utilized to assess the impact of selected variables on the tracking error of ETFs. The baseline model is defined as:

$$\mathrm{TE}_{i,t} = \alpha + \gamma \mathrm{TE}_{i,t-1} + \beta_1 \mathrm{TER}_{i,t} + \beta_2 \mathrm{AUM}_{i,t} + \beta_3 \mathrm{AGE}_{i,t} + \beta_4 \mathrm{RISK}_{i,t} + \beta_5 REP_{i,t} + \beta_6 ESG_{i,t} + u_{i,t} \tag{8}$$

where:

$TE_{i,t}$ – Tracking Error for ETF *i* at time *t*,

$TE_{i,t-1}$ – Tracking Error for ETF *i* at time *t-1*

α –the intercept representing the baseline level of TE when all other variables are zero,

$\gamma, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6$ – estimated coefficients for the respective variables,

TER – total expense ratio of an ETF *i* at time *t*,

AUM – assets under management of an ETF *i* at time *t*,

AGE – age of an ETF *i* at time *t*,

RISK – the underlying benchmark *i* risk at time *t*,

REP – dummy variable equal to 1 if an ETF uses a synthetic replication method,

ESG – dummy variable equal to 1 if an ETF tracks an ESG index,

u_{it} – the error term representing the residuals of the model.

To deepen the analysis, two separate models for ESG and non-ESG ETFs were estimated. Notably, the REP variable was not included in the analysis as only 2 ESG ETFs under study applied the synthetic replication. The model for each group is specified as follows:

$$\mathrm{TE}_{i,t} = \alpha + \gamma \mathrm{TE}_{i,t-1} + \beta_1 \mathrm{TER}_{i,t} + \beta_2 \mathrm{AUM}_{i,t} + \beta_3 \mathrm{AGE}_{i,t} + \beta_4 \mathrm{RISK}_{i,t} + u_{i,t} \tag{9}$$

where all variables remain as previously defined.

Finally, to further assess whether the determinants of tracking error differ significantly between ESG and non-ESG ETFs, an extended model incorporating interaction terms was estimated:

$$\begin{aligned}\mathrm{TE}_{i,t} = \alpha + \gamma \mathrm{TE}_{i,t-1} + \beta_1 \mathrm{TER}_{i,t} + \beta_2 \mathrm{AUM}_{i,t} + \beta_3 \mathrm{AGE}_{i,t} + \beta_4 \mathrm{RISK}_{i,t} + \beta_5 \mathrm{ESG}_{i,t} \\ + \beta_6(ESG \times \mathrm{TER}_{i,t}) + \beta_7(ESG \times \mathrm{AUM}_{i,t}) + \beta_8((ESG \times \mathrm{AGE}_{i,t)} \\ + \beta_9(ESG \times \mathrm{RISK}_{i,t}) + u_{i,t}\end{aligned} \tag{10}$$

where all variables remain as previously defined, with additional interaction terms capturing potential differences in the impact of TER, AUM, AGE, and RISK between ESG and non-ESG ETFs.

4. 3 Research Sample Characteristics

The final sample includes 48 passively managed ESG broad market ETFs listed on European exchanges and 86 comparable non-ESG ETFs. All funds are subject to similar European regulations and follow the same passive investment strategy. The study focuses on ETFs tracking global, regional, and country-specific indices, both ESG and non-ESG. The distribution shows a balanced mix across these geographical areas. However, in each category, non-ESG ETFs are more numerous than ESG ones, which reflects a general trend in the global market. As noted by Morningstar (2024), although the number of ESG funds has grown in recent years, they still represent a smaller share of the overall ETF market. Despite rising interest in sustainable investing, non-ESG ETFs continue to dominate. According to ETFGI (2024), as of February 24, 2024, there were around 1,406 ESG ETFs globally, compared to over 10,000 ETFs in total. In this study, ESG ETFs make up 35.55% of the sample, which is a strong representation considering their smaller share in the global market. Table 14 shows the breakdown by geographical focus.

Table 14. Regional exposure of ETFs in the research sample

Category of Equity Indices	Index Short Name	Number of ESG ETFs	Number of non-ESG ETFs	Total
Global Indices	FTSE Developed Index	1	2	33
	MSCI World Index	5	7	
	MSCI World Minimum Volatility Index	2	2	
	FTSE Emerging Index	1	2	
	MSCI Emerging Markets Index	3	8	
Regional Indices	MSCI Europe Index	8	9	50
	MSCI Europe Minimum Volatility Index	1	3	
	STOXX Europe 600 Index	3	6	
	MSCI EMU Index	5	9	
	MSCI Pacific ex Japan	1	5	
Country-specific Indices	MSCI Japan Index	4	6	51
	FTSE Japan Index	1	2	
	S&P 500 Index	4	12	
	MSCI USA Index	6	11	
	MSCI USA Small Cap Index	2	1	
	MSCI USA Minimum Volatility Index	1	1	
Total		48	86	134

Source: Own elaboration based on the Bloomberg database

Table 15 shows basic statistics for the assets under management (AUM) variable. Non-ESG ETFs have a higher average AUM, but ESG ETFs have a higher median value. The standard deviation is larger for non-ESG ETFs, which means higher variability in this group. The disparity in AUM is also highlighted by the range, where non-ESG ETFs have a minimum AUM of 3.83 USD million and a maximum of 17,699.75 USD million, contrasting with ESG ETFs' range of 10.69 USD million to 8,052.92 USD million.

Table 15. Descriptive statistics of ETF assets under management

Statistic	ESG ETFs (N=48)	Non-ESG ETFs (N=86)	Total (N=134)
Mean	1,108.39	1,678.81	1,474.48
Median	614.09	542.16	593.25
Standard Deviation	1,495.06	2,726.07	2,375.90
Minimum	10.69	3.83	3.83
Maximum	8,052.92	17,699.75	17,699.75

Note: Values in USD million. Data as of January 1, 2024.

Source: Own elaboration

The tracking error results may differ because of differences in fund size between ESG and non-ESG ETFs. Larger funds typically achieve lower costs and better liquidity, which results in reduced tracking error (Chu, 2011). The tracking errors of ESG ETFs in the sample appear more consistent but slightly higher because these funds have lower AUM and less variation, which will be analyzed in the study.

The data presented in Table 16 reveals that ESG ETFs exist for a shorter period than non-ESG ETFs. The average age of ESG ETFs amounts to approximately half of the average age of non-ESG ETFs. Non-ESG funds exist for a longer period and demonstrate greater age variability because they have established themselves in the market. ESG investing gained popularity recently because of increasing public understanding and supportive government regulations (Hartzmark and Sussman, 2019).

Table 16. Descriptive statistics of ETF age

Statistic	ESG ETFs (N=48)	Non-ESG ETFs (N=86)	Total (N=134)
Mean	5.06	10.55	8.59
Median	4.55	11.01	7.28
Standard Deviation	2.03	4.51	4.63
Minimum	3.34	3.26	3.26
Maximum	14.53	22.15	22.15

Note: Age expressed in years. Data as of January 1, 2024.

Source: Own elaboration

The comparison of total expense ratios (TER) between ESG and non-ESG ETFs reveals minimal differences in average and median values, with non-ESG ETFs having slightly higher costs (see Table 17). However, a notable distinction occurs in the variability of TERs. Non-ESG ETFs show a wider range and higher standard deviation of TER, which means their costs vary more.

Table 17. Descriptive statistics of ETF total expense ratios

Statistic	ESG ETFs (N=48)	Non-ESG ETFs (N=86)	Total (N=134)
Mean	0.19	0.20	0.19
Median	0.18	0.19	0.18
Standard Deviation	0.07	0.11	0.09
Minimum	0.07	0.05	0.05
Maximum	0.43	0.65	0.65

Note: Total expense ratio in %. Data as of January 1, 2024.

Source: Own elaboration

The lower TER in ESG ETFs may result from growing demand and stronger competition among fund managers trying to attract investors focused on sustainability. The National Bureau of Economic Research (NBER, 2024) points out that fund managers reduce fees to stay competitive and attract more capital inflows. Investors expect lower costs as part of their investment in sustainable products, and fund managers respond by adjusting their fee structures accordingly.

The Morningstar Sustainalytics (2024) data reveals that ESG funds in Europe have lower costs than traditional funds, and their costs have decreased substantially over the last ten years. The increasing demand for ESG investing has led to the creation of cost-efficient strategies among other factors. Then, the simplification of the ESG criteria has reduced costs for ESG analysis, data collection, and engagement activities, which subsequently decreases the total expense ratio (TER).

The lower variability in TER for ESG ETFs may indicate that these funds have more standardized fee structures, possibly due to regulatory pressures and a focus on transparency and accountability. For example, the PwC Global Investment (2022) survey showed that investors value financial discipline and greater transparency in sustainability reports to trust the information provided by companies. The increased regulatory scrutiny and demands for greater transparency in ESG practices put pressure on managers to offer cost-effective solutions.

Both ESG and non-ESG ETFs commonly adopt physical replication techniques (see Table 18). 95.83% of ESG ETFs implement full or optimized replication methods. Synthetic replication appears more often in non-ESG ETFs because 22 out of 86 funds use derivatives to track their indices. The extensive adoption of physical replication by ESG ETFs may lead to higher tracking errors, as suggested by previous research by Elia (2012) and Johnson et al. (2013).

Table 18. Distribution of ETF replication methods

Replication Method	ESG ETFs (N=48)	Non-ESG ETFs (N=86)	Total (N=134)
Physical	46	64	110
Synthetic	2	22	24

Note: Data as of January 1, 2024.

Source: Own elaboration

Table 19 shows that ESG ETFs maintain benchmark risk levels comparable to non-ESG ETFs. The slightly increased maximum risk in ESG ETFs indicates they might exhibit additional risk under specific circumstances. These findings align with previous studies by Friede et al. (2015) and Revelli & Viviani (2015), who proved that ESG investments have comparable risk profiles to traditional investments.

Table 19. Descriptive statistics of benchmark risk

Statistic	ESG ETFs (N=48)	Non-ESG ETFs (N=86)	Total (N=134)
Mean	0.02	0.02	0.02
Median	0.02	0.02	0.02
Standard Deviation	0.00	0.00	0.00
Minimum	0.02	0.02	0.02
Maximum	0.03	0.02	0.03

Note: Benchmark risk expressed as % calculated over the full research period.

Source: Own elaboration

The following analysis requires verification of normal distribution and equality of variances between ESG and non-ESG ETFs. The Shapiro–Wilk test and Levene’s test results appear in Table 20. The Shapiro–Wilk test results demonstrate that all independent variables fail to meet normal distribution criteria because their p-values remain below 0.05. The results of Levene’s test indicate that AGE and RISK demonstrate unequal variances because their p-values are less than 0.05, but TER and AUM show equal variances since their p-values exceed 0.05.

Table 20. Results of normality and homogeneity tests for independent variables

Test	Shapiro-Wilk Test		Levene's Test	
Variable	**W Statistic**	**p-value**	**F Statistic**	**p-value**
TER	0.86	0.00	3.72	0.06
AUM	0.25	0.00	1.58	0.21
AGE	0.89	0.00	34.39	0.00
RISK	0.95	0.00	4.42	0.04

Source: Own elaboration

Given the violation of normality and homogeneity, the non-parametric Wilcoxon rank sum test is applied to compare the distributions of independent variables for ESG and non-ESG ETFs (see Table 4.11).

Table 21. Significance of differences in independent variables

Test	**Wilcoxon rank sum test**	
Variable	**W Statistic**	**p-value**
TER	2237	0.96
AUM	2333	0.72
AGE	3969.5	0.00
RISK	1791	0.04

Source: Own elaboration

The high p-values for TER and AUM indicate that there are no significant differences between ESG and non-ESG ETFs in terms of expenses and size. On the other hand, low p-values for AGE and RISK indicate that there are significant differences in maturity and benchmark risk between the two fund groups.

Figure 15 shows that the TER distributions for ESG and non-ESG ETFs are similar. In both cases, the distributions are right-skewed and concentrated between 0.1% and 0.3%. ESG ETFs show slightly lower variability, which means that the fee structures are more uniform.

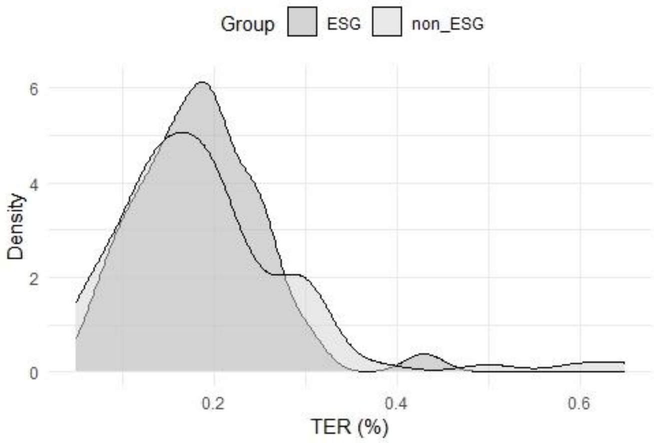

Figure 15. Distribution of total expense ratio (TER) for ESG and Non-ESG ETFs (Source: Own elaboration)

Then, Figure 16 shows the similarity of the AUM distributions between ESG and non-ESG funds. The AUM distributions are highly right-skewed in both cases, reflecting a predominance of smaller funds in the sample. However, a smaller share of ESG ETFs exceeds the 2,000 million USD threshold (16.66% of funds) compared to non-ESG ETFs (27.91% of funds), indicating that despite their rapid growth, ESG ETFs have not yet achieved the same scale as their non-ESG counterparts.

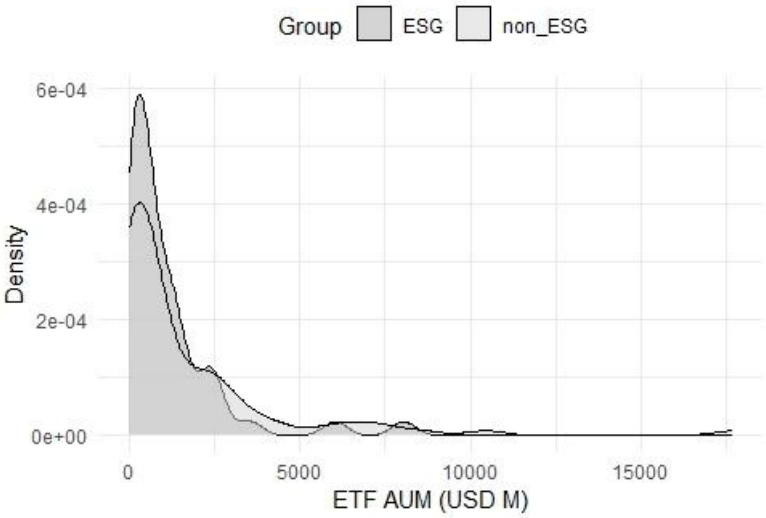

Figure 16. Distribution of assets under management (AUM) for ESG and Non-ESG ETFs (Source: Own elaboration)

Figure 17 shows clear differences in the AGE distribution between ESG and non-ESG ETFs. Non-ESG ETFs display multiple peaks, suggesting a mix of age groups. On the other hand, ESG ETFs are more concentrated regarding maturing, with a strong peak at 3–4 years, reflecting their shorter market presence.

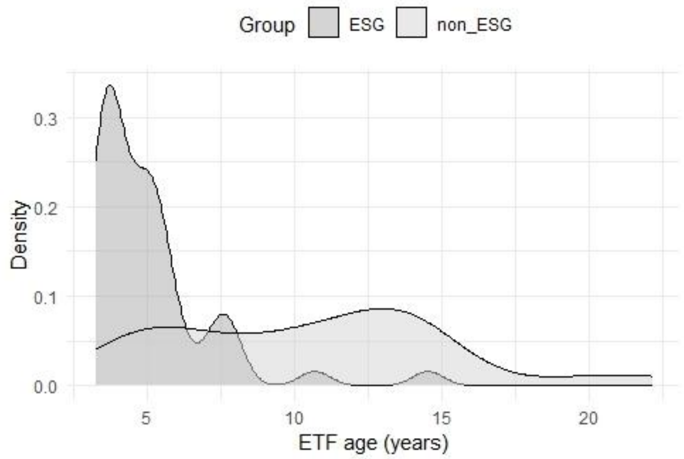

Figure 17. Distribution of age for ESG and Non-ESG ETFs (Source: Own elaboration)

Figure 18 depicts the distribution of the RISK variable, highlighting differences between ETF categories. The risk profiles of ESG ETFs show multiple peaks in their density plot. Then, non-ESG ETFs show a more focused distribution of risk. The difference in risk distribution is probably because many non-ESG ETFs track the same index, while each ESG ETF tracks its own ESG index.

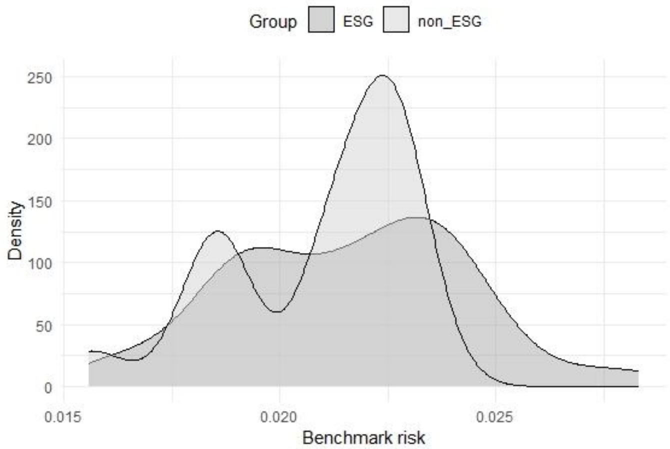

Figure 18. Distribution of benchmark risk (RISK) for ESG and Non-ESG ETFs
(Source: Own elaboration)

The final part of independent variable exploration involves correlation analysis. The correlation heatmap presented in Figure 19 shows that the investigated variables do not show a significant correlation. The strongest relationship is a negative correlation (-0.21) between Total Expense Ratio (TER) and Assets Under Management (AUM), indicating that larger ETFs tend to have lower expense ratios.

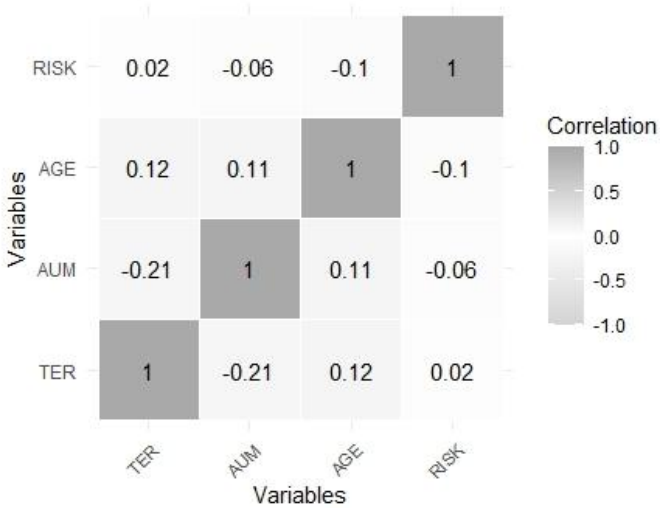

Figure 19. Correlation heatmap of variables under study
(Source: Own elaboration)

All in all, based on the analysis of ETF features, the tracking errors of ESG ETFs might be higher than those of non-ESG ETFs because of their lower maturity and frequent use of physical replication. On the other hand, the total expense ratios appear to be comparable in the two groups. The next part of the paper will thoroughly examine and compare the tracking errors of ESG and non-ESG ETFs to provide a comprehensive understanding of their tracking ability.

4. 4 Empirical Results and Discussion

4. 4. 1 Comparison of Tracking Errors in ESG and Non-ESG ETFs

The evaluation of tracking errors between ESG and non-ESG ETFs confirms that ETFs track their benchmarks with exceptionally high precision. The tracking errors of 130 out of 134 observed ETFs remain below 0.5%, which Banerjee (2015) established as the internationally accepted threshold. None of the ETFs under investigation reached the tracking error beyond 0.53%. As detailed in Table 22, ESG ETFs generally demonstrate slightly lower tracking errors compared to their non-ESG counterparts. This trend is consistent across different tracking error measures, such as TE_1, TE_2, and TE_3.

The tracking errors exhibit higher variation between ESG and non-ESG ETFs for TE_1 and TE_3 measures than the TE_2 measure. This is the TE_2 that shows the standard deviation of return differences. The TE_1 measure shows the direct return difference between the ETF and its benchmark, while TE_3 measures the degree of return deviation from market exposure expectations. This discrepancy suggests that ESG and non-ESG ETFs exhibit greater similarity in the consistency of their tracking performance while showing slightly higher differences in the magnitude of deviation from their benchmarks. Subsequent steps in the research involve verifying the significance of these differences.

Table 22. Descriptive statistics of tracking errors (%) for ESG and non-ESG ETFs

Variable	Mean	Median	Std Dev	Min	Max
ESG ETFs					
TE_1	0.11	0.08	0.09	0.01	0.44
TE_2	0.04	0.03	0.03	0.00	0.11
TE_3	0.11	0.08	0.09	0.01	0.43
Non-ESG ETFs					
TE_1	0.16	0.12	0.15	0.00	0.53
TE_2	0.06	0.04	0.09	0.00	0.39
TE_3	0.16	0.12	0.14	0.00	0.52

Source: Own elaboration

The tracking error measures show significant correlations with each other, as shown in Figure 20. High correlation values (above 0.8) indicate strong similarity between the measures, so they may be used interchangeably.

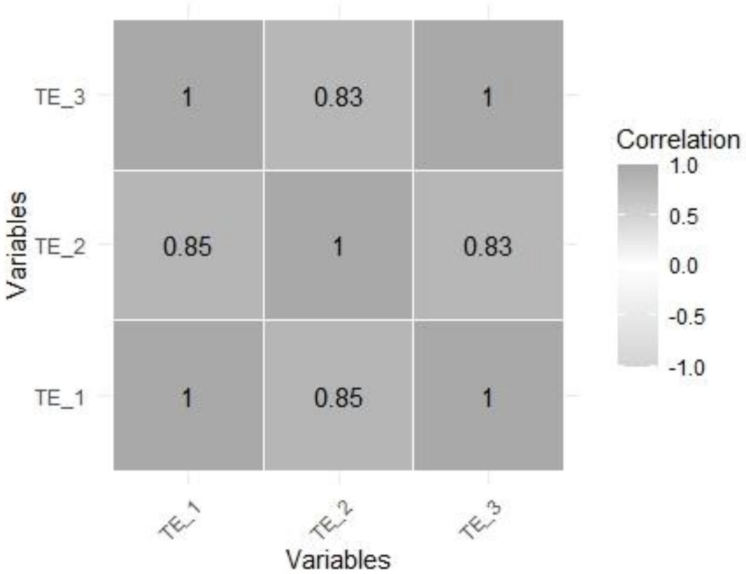

Figure 20. Correlation heatmap of ETFs' tracking errors
(Source: Own elaboration)

The Shapiro-Wilk and Levene's tests (see Table 23) show that there are violations of normality and homogeneity of variances for all dependent variables ($p < 0.05$). Therefore, the Wilcoxon rank-sum test is used to compare the tracking error distributions between groups.

Table 23. Normality and homogeneity assumptions for dependent variables

Test	**Shapiro-Wilk Test**		**Levene's Test**	
Variable	**W Statistic**	**p-value**	**F Statistic**	**p-value**
TE_1	0.86	0.00	8.40	0.00
TE_2	0.60	0.00	3.99	0.04
TE_3	0.87	0.00	8.46	0.00

Source: Own elaboration

The Wilcoxon rank-sum test results appear in Table 24. The null hypothesis about equal tracking error distributions between ESG and non-ESG ETFs is rejected as p-values exceed the 0.05 significance threshold. The results show that previously observed tracking error differences are statistically insignificant.

Table 24. Significance of differences between tracking errors of ESG and Non-ESG ETFs

Test	Wilcoxon rank sum test	
Variable	**W Statistic**	**p-value**
TE_1	1720	0.09
TE_2	1820	0.22
TE_3	1680	0.08

Source: Own elaboration

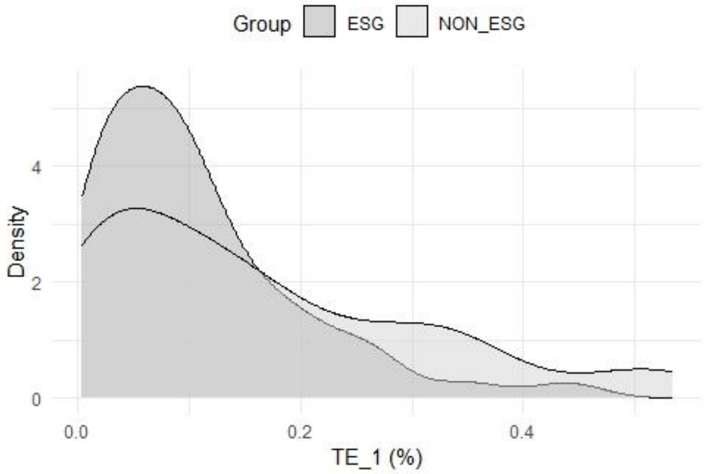

Figure 21. Distribution of tracking error (TE_1) for ESG and Non-ESG ETFs (Source: Own elaboration)

As shown in Figure 21, both groups exhibit similar patterns of the TE_1 variable distribution. However, TE_1 values are more tightly clustered below 0.2% for ESG ETFs, while in the case of non-ESG ETFs, there are more tracking errors in the 0.2%–0.5% range.

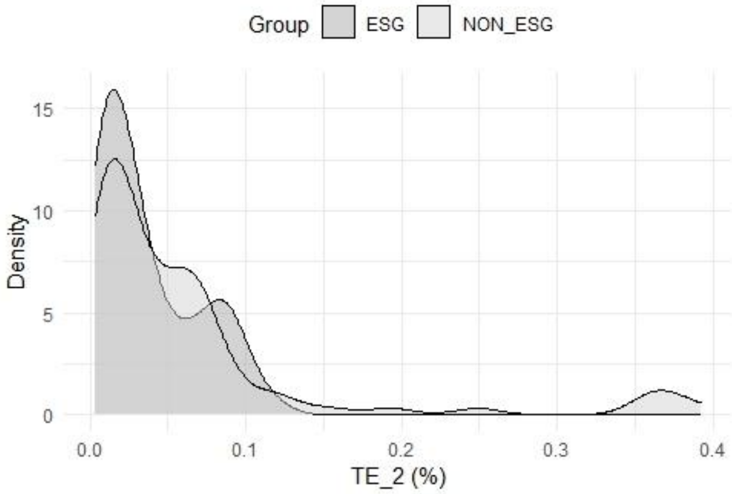

Figure 22. Distribution of tracking error (TE_2) for ESG and Non-ESG ETFs (Source: Own elaboration)

The tracking error distribution of the TE_2 variable (see Figure 22) shows a high degree of similarity for ESG and non-ESG ETFs. Both groups exhibit concentrated distributions below 0.1% with peak densities in the 0%–0.05% range. A long right tail, occurring both for ESG and non-ESG ETF distributions, indicates the magnitude of higher tracking errors.

The peak density for both ESG and non-ESG ETFs occurs at similar TE_3 values (see Figure 23), from around 0.05% to 0.1%. This indicates that both types of ETFs have a similar concentration of funds with low tracking errors. Then, two groups of funds have a long tail extending towards higher TE_3 values (up to 0.5%). This means that while some ETFs in both categories have higher tracking errors, these instances are relatively rare.

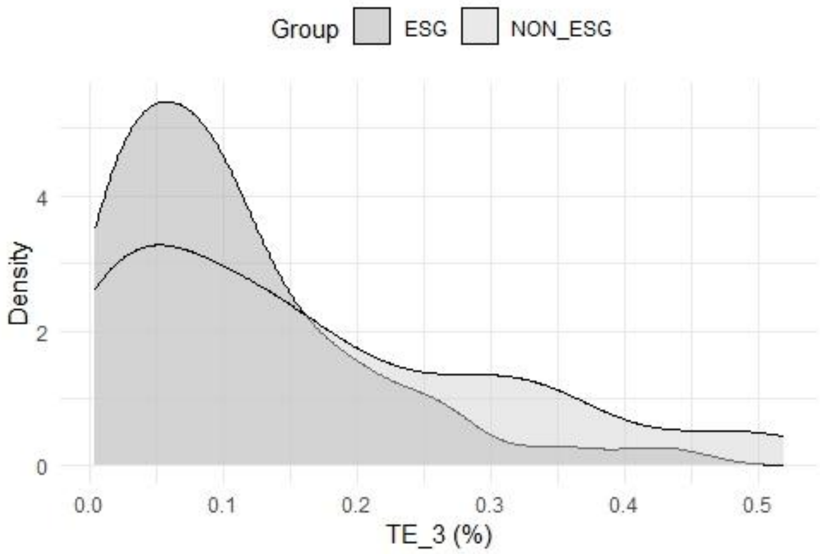

Figure 23. Distribution of tracking error (TE_3) for ESG and Non-ESG ETFs
(Source: Own elaboration)

The density plots of TE_1, TE_2, and TE_3 variables for ESG and non-ESG ETFs show that both categories of ETFs have similar tracking abilities with only slight differences. The results indicate that the tracking ability of passive ESG equity ETFs listed on European exchanges is statistically indistinguishable from that of their non-ESG counterparts. However, these findings are based on cross-sectional data and do not account for the time component. Therefore, in the subsequent section of the study, the differences in ETFs' tracking performance will be verified using panel data depicting both numerous entities and time series.

4. 4. 2 Modelling Tracking Error of ETFs

The following section describes the implementation of the baseline model according to Equation (8). The analysis in this part differs from the previous study because it uses panel data from 134 ETFs across 42 periods, which generates 5,628 observations. Descriptive statistics for the scaled continuous variables used in the model are shown in Table 25. The results indicate that most variables are concentrated in the lower part of the [0,1] range. The distribution of TE_2 is highly left-skewed, as indicated by a low mean and median. This confirms that most ETFs exhibit minimal tracking errors. Similarly, the distribution of AUM is left-skewed, reflecting a predominance of smaller funds in the sample. Then, TER, AGE, and RISK display slightly higher dispersion, indicating greater variability across observations.

Table 25. Descriptive statistics of variables used in the model (full ETF sample)

Variable	Mean	Median	Std Dev	Min	Max	N
TE_2	0.03	0.00	0.07	0.00	1.00	5,628
TER	0.24	0.22	0.15	0.00	1.00	5,628
AUM	0.05	0.02	0.09	0.00	1.00	5,628
AGE	0.18	0.16	0.11	0.00	1.00	5,628
RISK	0.24	0.20	0.14	0.00	1.00	5,628

Source: Own elaboration

Initially, the study uses correlation analysis to evaluate the relationships between variables. The non-normal distribution of data and the presence of binary variables led to the selection of Spearman’s method. The Spearman correlation coefficients between the investigated variables appear in Table 26.

Table 26. Correlation matrix for the full sample of ETFs

	TE_2	TER	AUM	RISK	AGE	ESG	REP
TE_2	1*						
TER	0.11*	1*					
AUM	-0.04*	-0.28*	1*				
RISK	0.03*	0	0.01	1*			
AGE	0.05*	0.08*	0.14*	0.05*	1*		
ESG	-0.05*	0.02	-0.09*	0	0.49*	1*	
REP	-0.15*	-0.07*	0.05*	0	-0.15*	-0.25*	1*

* Correlation is significant at the 0.05 level (two-tailed).

Source: Own elaboration

The results suggest that TE_2 increases with higher fees (TER), greater benchmark risk, and fund age, meaning that older and riskier ETFs with higher costs tend to have larger tracking errors. In contrast, ETFs with more assets (AUM), those following ESG indices, and those using synthetic replication tend to have lower tracking errors. All independent variables show weak correlations with TE_2 (between -0.15 and 0.11) and are statistically significant at the 0.05 level. Importantly, the direction of these relationships is consistent with earlier research on ETF tracking error (Chu, 2011; Rompotis, 2011), which provides a solid foundation for further econometric analysis.

To visualize the relationships between the independent variables and TE_2, a scatter plot analysis was conducted (see Figure 24). The plot's highly dispersed points indicate that the interactions between variables are not identifiable, suggesting no strong associations. The plot AUM vs TE_2 implies that larger ETFs tend to display lower tracking errors, however, this association is weakened by the high data volatility. The second plot (TER vs TE_2) and the third plot (TE_2 vs RISK) do not show any clear trends, which means that these variables have little impact on TE_2. Then, the last plot (TE_2 vs AGE) shows that younger ETFs tend to have higher tracking errors. In general, these results are in line with the earlier correlation analysis.

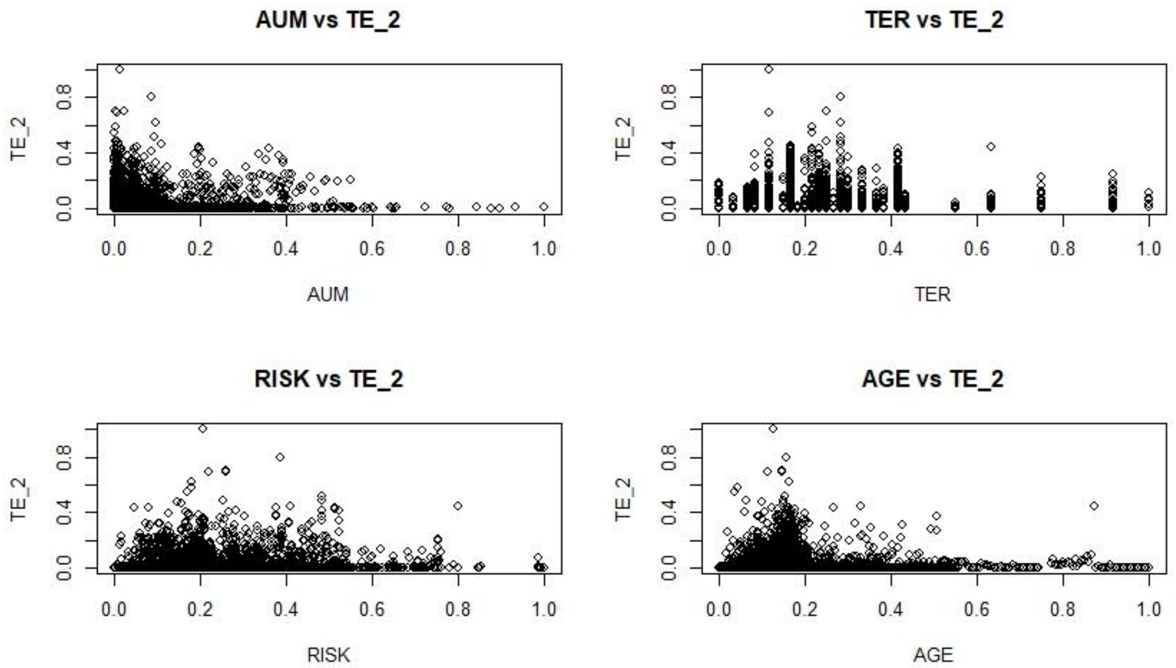

Figure 24. Dependencies between variables for the full sample of ETFs
(Source: Own elaboration)

Then, the Levin-Lin-Chu unit-root test was employed to assess the stationarity of the variables (Levin et al., 2002). Since the TER variable was assumed to be constant throughout the entire period for each ETF, it is inherently stationary and did not require a unit-root test. The results presented in Table 27 indicate that TE_2 and RISK variables are stationary. However, the AUM and AGE variables exhibited the presence of unit roots.

Table 27. Results of the Levin–Lin–Chu unit root test (full sample)

Variable	Z Statistic	p-value
TE_2	-16.85	0.00
TER	–	–
AUM	10.80	1.00
RISK	-13.50	0.00
AGE	7.34	1.00

(Source: Own elaboration)

To address the problem of non-stationarity, first-order differencing was used. This means that for each value, the previous value was subtracted. The ADF test (see Table 28) confirmed that all variables became stationary after this step.

Table 28. Results of the ADF test (full sample)

Variable	ADF Statistic	p-value
TE_2	-20.49	0.01
TER	–	–
AUM	-16.75	0.01
RISK	-25.21	0.01
AGE	-9.23	0.01

(Source: Own elaboration)

Next, static panel models, including Pooled OLS, Random Effects, and Fixed Effects, were used to explore direct relationships between tracking error and selected predictors, without including lagged effects (see Table 29).

Table 29. Comparison of Pooled OLS, RE, and FE models (full sample)

Model Formula	TE_2 ~ TER + AUM + RISK + AGE + ESG + REP					
Variable	**Pooled OLS**		**Random Effects (RE)**		**Fixed Effects (FE)**	
	estimate	p-value	estimate	p-value	estimate	p-value
TER	0.03	0.00	0.02	0.28	–	–
AUM	0.03	0.82	-0.16	0.14	-0.18	0.12
RISK	0.02	0.00	0.01	0.01	0.01	0.02
AGE	-3.35	0.00	-0.88	0.43	-0.70	0.53
ESG	-0.01	0.32	-0.02	0.05	–	–
REP	-0.03	0.00	-0.03	0.00	–	–
Intercept	0.02	0.00	0.03	0.00	–	–
R-squred	0.0424		0.0046		0.0016	
Adj *R*-squred	0.0413		0.0035		-0.0236	

(Source: Own elaboration)

As shown in Table 29, model results differ in both coefficient significance and direction. Low R-squared values suggest that important factors affecting TE_2 may be missing or not well captured. The Pooled OLS model found TER, RISK, AGE, and REP as significant predictors of TE_2. However, the model explained only 4.2% of its variation, which indicates extremely limited explanatory power. Furthermore, the significant intercept indicates that there are omitted variables. In the Random Effects model, only RISK and REP factors remained significant. However, the very low R-squared value indicates that accounting for differences between ETFs does not improve model performance. The intercept again suggests omitted factors (Greene, 2018: 87). The Fixed Effects model, which controls for ETF-specific characteristics, found only the RISK variable to be significant. A negative adjusted R-squared shows poor model fit, which questions the reliability of the results.
To sum up, the results indicate the need for a deeper analysis of the data to identify the reasons behind the model's poor fit. Therefore, the next step is to select the best static model (Pooled OLS, Fixed Effects, Random Effects), and analyse its properties in detail. To do this, several diagnostic tests were carried out, as shown in Table 30.

Table 30. Diagnostic tests for model selection

Test	**Statistic**	**Degrees of Freedom**	**p-value**
F-test (Pooled vs FE)	F = 20.65	Df1 = 130, Df2 = 5357	0.00
LM test (Random vs Pooled)	χ^2 = 10952	df = 1	0.00
Hausman (RE vs FE)	χ^2 = 21.06	Df1 = 130, Df2 = 5491	0.00

(Source: Own elaboration)

The F-test rejects the null hypothesis of no fixed effects, indicating that the Fixed Effects (FE) model fits better than the Pooled OLS model. The LM test also rejects the null of no random effects, favouring the Random Effects (RE) model over Pooled OLS. Finally, the Hausman test shows that the FE model is more appropriate than the RE model. Overall, these tests confirm that the Fixed Effects model best explains the tracking error of the ETFs in this analysis.

The diagnostic results in Table 31 reveal key weaknesses of the Fixed Effects model. Both the Wooldridge and Breusch-Godfrey tests show that residuals are autocorrelated, meaning that tracking errors are linked over time. The Wooldridge test detects first-order autocorrelation, while the Breusch-Godfrey test finds higher-order patterns. This confirms that time dependence exists in the data, making static models unsuitable for analyzing ETF tracking errors.

Table 31. Diagnostic tests for the Fixed Effects model (full ETF sample)

Test	Statistic	p-value
Wooldridge Test	$\chi^2 = 845.56$	0.00
Breusch-Godfrey Test	LM = 792.38	0.00
Breusch-Pagan Test	BP = 67.23	0.00
Jarque-Bera Test	X-squared = 437182	0.00
RESET Test	RESET = 1.63	0.20

(Source: Own elaboration)

Figure 25 shows that past values of TE_2 impact its current values. Even though the effect diminishes with more lags, it is still important. Previous studies, like DeFusco et al. (2011) and Ivanov (2015), also found that tracking errors in ETFs follow patterns that can be captured by autoregressive models. This implies that a dynamic model may be better for explaining changes in tracking error.

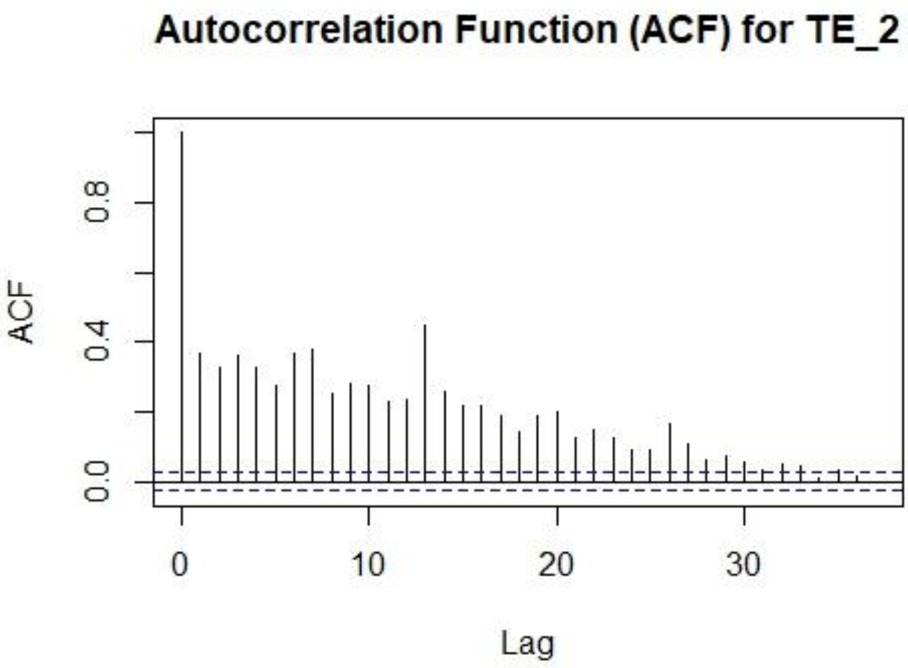

Figure 25. Autocorrelation function (ACF) for TE_2 for the full sample of ETFs
(Source: Own elaboration)

The Breusch-Pagan test (see Table 31) shows that the model suffers from heteroskedasticity, meaning the error variance is not constant. To correct this, robust standard errors should be used (White, 1980). Moreover, the Jarque-Bera test indicates that the residuals are not normally distributed, which reduces the reliability of the results. Although the RESET test confirms that the model is correctly specified in terms of functional form and included variables, it does not detect problems like heteroskedasticity or non-normality. These issues suggest the need for further model improvements or alternative methods.

The final diagnostic tests for the Fixed Effects model, including multicollinearity and endogeneity tests, are presented in Table 32. The Variance Inflation Factor (VIF) was used to assess multicollinearity. Most variables have VIF values well below 10, indicating no concern of multicollinearity. However, AGE and ESG are close to the threshold, suggesting a moderate correlation. Since ESG ETFs are generally younger, this reflects a real pattern in the data rather than a modeling flaw. To address this, separate models for ESG and non-ESG ETFs should be estimated, removing the ESG variable and reducing multicollinearity.

The Durbin-Wu-Hausman (DWH) test suggests possible endogeneity, as low p-values indicate correlation between some regressors and the error term. This may be linked to the earlier-detected autocorrelation, suggesting that omitted lagged effects bias the estimates. To address this, including a lagged dependent variable and applying a dynamic panel model, such as the GMM estimator proposed by Arellano and Bond (1991), is advised. GMM effectively deals with both endogeneity and autocorrelation by using internal instruments (Wooldridge, 2010; Baum et al., 2003).

Table 32. Tests for endogeneity and multicollinearity in the Fixed Effects model (full sample)

Explanatory Variable	**VIF**	**DWH Statistic**	**DWH p-value**
TER	1.02	6.95	0.01
AUM	1.02	4.84	0.03
RISK	1.00	13.00	0.00
AGE	9.51	6.89	0.01
ESG	9.64	6.95	0.01
REP	1.07	6.95	0.01

(Source: Own elaboration)

In conclusion, the Fixed Effects, Random Effects, and Pooled OLS models fail to model ETF tracking errors because they lack explanatory power and produce diagnostic issues, including autocorrelation, heteroskedasticity, and endogeneity. The FE model controls for time-invariant differences, but it does not capture the dynamic relationships. The dynamic panel GMM estimator proposed by Arellano and Bond (1991) provides a more robust solution. The model details are presented in Table 33.

Table 33. Specification of the linear dynamic panel model (full ETF sample)

Specification	Description
Lagged Dependent Variable	$TE\_2_{i,t-1}$ (First-order lag of tracking error)
Instrumental Variables	lag(TE_2, 20) (Twenty-period lag of tracking error)
Unit Effects	Individual fixed effects
Estimation Approach	One-step Generalized Method of Moments (GMM), Arellano-Bond
Data Transformation	First-difference transformation (Δ)
Robust Standard Errors	Heteroskedasticity and autocorrelation-consistent (HAC) robust standard errors

(Source: Own elaboration)

The dynamic model includes the first lag of TE_2 to capture time-related effects and addresses the problem of autocorrelation. The GMM estimator is used to address endogeneity, following Arellano and Bond (1991) and Blundell and Bond (1998), with internal instruments based on lagged dependent variables. A set of lagged levels of TE_2 was tested as instruments (see Table 34). Lag(TE_2, 20) offered the best trade-off between relevance and exogeneity based on the Hansen J-test and absence of second-order autocorrelation (AR(2)). Although some lags showed higher correlations with TE_2, they failed validity tests.

Table 34. Evaluation of lagged TE_2 as an instrumental variable in the GMM model (full sample)

Lag TE_2	Correlation with TE_2	AR (2) Test Statistic	AR (2) p-value	Hansen J-Test Statistic	Hansen p-value
2	0.34	0.22	0.82	118.88	0.01
3	0.37	-0.94	0.35	126.19	0.00
4	0.34	-0.15	0.88	122.38	0.00
5	0.29	0.78	0.44	121.99	0.01
6	0.38	0.18	0.86	124.31	0.00
7	0.40	-1.07	0.29	126.97	0.00
8	0.26	-0.14	0.89	122.09	0.01
9	0.29	-0.66	0.51	124.45	0.00
10	0.29	0.90	0.37	119.41	0.01
11	0.24	-0.78	0.44	121.04	0.01
12	0.24	-1.95	0.05	121.64	0.01
13	0.45	0.22	0.82	119.36	0.01
14	0.27	-2.88	0.00	126.52	0.01
15	0.23	-1.46	0.14	116.12	0.03
16	0.23	-0.26	0.80	116.28	0.04
17	0.20	0.06	0.95	107.62	0.13
18	0.15	-0.12	0.90	111.54	0.09
19	0.20	-1.22	0.22	106.03	0.21
20	0.21	-0.17	0.87	105.77	0.25

(Source: Own elaboration)

Next, the model addresses individual effects, i.e., time-invariant differences between entities (ETFs). This permits controlling for specific characteristics of ETFs, eliminating the influence of such constant factors. Finally, as the raw data was not stationary, the first-difference transformation was applied. This method involves transforming the data by subtraction the previous period's variable values from the current values. Consequently, instead of focusing on the formation of absolute TE_2 values, the model depicts changes in tracking error over time.

The model was estimated using the one-step GMM estimator. According to Arellano and Bond (1991), this method helps prevent overfitting, making the results more stable and interpretable. Robust HAC standard errors were applied to improve accuracy by correcting for heteroskedasticity and autocorrelation. The estimation results are in Table 35. Importantly, as the model applies first-difference transformations, estimates should be interpreted in terms of changes between periods rather than levels of variables (Baltagi, 2015: 161). The obtained estimates reflect the effect of changes in the explanatory variables on changes in the tracking error.

Table 35. Estimation results of the GMM model (full ETF sample)

Model Formula	TE_2 ~ lag (TE_2, 1) + TER + AUM + RISK + AGE + ESG + REP | lag (TE_2,20)			
Variable	**Estimate**	**Std error**	**Statistic**	**p-value**
lag (TE_2, 1)	0.29	0.06	5.25	0.00
TER	0.04	0.01	3.80	0.00
AUM	0.06	0.04	1.57	0.12
RISK	0.03	0.01	4.40	0.00
AGE	0.02	0.02	1.18	0.24
ESG	-0.01	0.00	-3.00	0.00
REP	-0.02	0.00	-3.84	0.00

(Source: Own elaboration)

The estimate for the lagged tracking error (lag_TE_2) is statistically significant. It evidences that tracking error changes tend to continue from one period to the next. The autoregressive pattern of the variable becomes evident because tracking error decreases in one period tend to lead to decreases in the following period. Next, the model confirms that the total expense ratio (TER) has a significant relationship with tracking error. It shows that tracking error increases with an increase in TER. The low standard error value demonstrates that TER provides precise tracking performance predictions. The model also shows that assets under management (AUM) do not have a significant impact on the tracking error of ETFs.

The insignificant coefficient, together with its high standard error value, indicates that fund size has no relevant effect on tracking errors.

Another variable displaying a significant impact on the tracking error of ETFs is benchmark risk (RISK). The estimate indicates that an increase in benchmark risk from one period to the next increases the tracking error of an ETF. Then, the low estimate for the AGE variable and the relatively high standard error imply that the fund age does not have a significant impact on its tracking error. The age of the ETF is therefore not a decisive determinant of the TE_2 variable, implying that ETFs with different market experiences have similar abilities in accurately replicating the index.

A statistically significant negative estimate for the ESG variable means that ETFs tracking ESG indices systematically exhibit lower tracking errors than non-ESG funds. The value of the estimate (-0.01) indicates that these are not large differences, however, they are noticeable. The significant impact of the ESG variable in the model implies the confirmation of the first research hypothesis (H1), stating that passive ESG equity ETFs listed on European exchanges exhibit significantly different tracking errors compared to their non-ESG counterparts.

For the variable REP, corresponding to ETFs using synthetic replication, the coefficient of the estimate is statistically significant. Although the estimate is relatively low, its precision (low standard error) suggests that ETFs using synthetic replication systematically have lower tracking errors than ETFs using physical replication.

Table 36. Diagnostic tests of the GMM model (full ETF sample)

Test	**Statistic**	**p-value**
Autocorrelation Test AR(1)	-5.96	0.00
Autocorrelation Test AR(2)	-0.17	0.87
Hansen Test	$\chi^2 = 105.77$	0.25
Wald Test	$\chi^2 = 101.63$	0.00

(Source: Own elaboration)

Table 36 shows the results of the model’s diagnostic tests. The AR(1) test confirms first-order autocorrelation in the residuals, which is typical for dynamic panel models. The AR(2) test does not show second-order autocorrelation, which means the model’s errors are not linked to their earlier values. The high p-value in the Hansen test means the instruments used in the model are correct and not related to the errors. This reinforces the estimation’s reliability and stability. Lastly, the Wald test points to the joint significance of the model parameters, confirming that the explanatory variables significantly account for the variation in tracking error. All in all, the results imply the model's strong fit and the relevance of the selected variables in explaining changes in tracking error.

The robustness check is the final step of the analysis. In this part, the model was re-estimated using a different tracking error measure (see Table 37). Specifically, TE_1, which depicts the absolute difference between fund and index returns, was used as a dependent variable. The results are consistent with those in Table 35. The key variables influencing tracking error, as well as the direction of their impact, remained the same when switching from TE_2 to TE_1. Although some coefficient values changed slightly, these differences do not affect the overall conclusions of the study.

Table 37. Robustness checks for the GMM model (full ETF sample)

Model Formula	TE_1 ~ lag (TE_1, 1) + TER + AUM + RISK + AGE + ESG + REP | lag(TE_1,20)			
Variable	**Estimate**	**Std error**	**Statistic**	**p-value**
lag(TE_1, 1)	0.20	0.04	4.92	0.00
TER	0.06	0.02	3.48	0.00
AUM	0.07	0.04	1.57	0.12
RISK	0.06	0.01	3.91	0.00
AGE	0.08	0.05	1.74	0.08
ESG	-0.04	0.01	-4.70	0.00
REP	-0.02	0.01	-2.24	0.03

(Source: Own elaboration)

4. 4. 3 Tracking Error Determinants in ESG and Non-ESG ETFs

The significance of the ESG variable in the model depicting the tracking error of ETFs suggests that ESG ETFs exhibit unique features regarding their tracking ability. This section presents the development of two separate models for ESG and non-ESG ETFs specified in Equation (9).

Table 38. Descriptive statistics of continuous variables for ESG and non-ESG ETFs

Variable	**ETF Group**	**Mean**	**Median**	**Minimum**	**Maximum**	**Std. Dev.**
TE_2	Non-ESG	0.03	0.01	0.00	1.00	0.08
	ESG	0.03	0.00	0.00	1.00	0.08
TER	Non-ESG	0.32	0.31	0.00	1.00	0.19
	ESG	0.24	0.23	0.00	1.00	0.17
AUM	Non-ESG	0.09	0.04	0.00	1.00	0.15
	ESG	0.06	0.02	0.00	1.00	0.10
RISK	Non-ESG	0.24	0.20	0.00	1.00	0.15
	ESG	0.24	0.19	0.00	1.00	0.15
AGE	Non-ESG	0.24	0.21	0.00	1.00	0.16
	ESG	0.66	0.72	0.00	1.00	0.20

(Source: Own elaboration)

The descriptive statistics presented in Table 38 demonstrate that ESG and non-ESG ETFs maintain comparable average tracking errors. The average TER, AUM, and age of non-ESG ETFs exceed that of ESG ETFs. These distinct structural features of these groups justify separate modeling approaches. The ADF test results presented in Table 39 demonstrate that all variables achieve stationarity after applying first differences.

Table 39. ADF test results for ESG and non-ESG ETFs

Variable	ESG ETFs		Non-ESG ETFs	
	ADF Statistic	**p-value**	**ADF Statistic**	**p-value**
TE_2	-14.71	0.01	-21.91	0.01
TER	–	–	–	–
AUM	-10.71	0.01	-12.74	0.01
RISK	-19.06	0.01	-26.75	0.01
AGE	-16.57	0.01	-10.70	0.01

(Source: Own elaboration)

At first, the regression analysis included adopting a static panel approach, like Pooled OLS, Random Effects (RE), and Fixed Effects (FE) models. However, the extremely low R-squared values showed that, similarly to the overall population of ETFs, these simple classical regression models do not perform well in describing the tracking error of ESG and non-ESG ETFs separately.

The diagnostic tests in Table 40 indicate that Fixed Effects models are not justified in the subgroup analyses, likely due to reduced heterogeneity within ESG and non-ESG ETFs. The lack of significant differences between FE and RE estimators indicated by the Hausman test suggests that unobserved effects are not correlated with regressors. This suggests that in the more homogeneous ESG and non-ESG subgroups, RE provides consistent and more efficient estimates.

Table 40. Diagnostic tests for model selection in ESG and non-ESG ETF samples

Test	ESG ETFs		Non-ESG ETFs	
	Statistic	**p-value**	**Statistic**	**p-value**
F-test (Pooled vs FE)	F = 0.12	1.00	F = 0.01	1.00
LM test (Random vs Pooled)	$\chi^2 = 19.21$	0.00	$\chi^2 = 43.50$	0.00
Hausman (RE vs FE)	$\chi^2 = 0.04$	1.00	$\chi^2 = 0.12$	0.99

(Source: Own elaboration)

Table 41 shows the diagnostic tests for the Random Effects models for ESG and non-ESG ETFs. The Breusch-Godfrey test confirms serial correlation in both models. The Breusch-

Pagan test finds no heteroskedasticity, so the residuals have stable variance. The Jarque-Bera test rejects normality of residuals, which may affect standard inference. The RESET test suggests that the model's functional form is correctly specified.

Table 41. Diagnostic tests for Random Effects models in ESG and non-ESG ETFs

Test	**ESG ETFs**		**Non-ESG ETFs**	
	Statistic	**p-value**	**Statistic**	**p-value**
Breusch-Godfrey Test	χ^2 = 527.24	0.00	χ^2 = 1958.1	0.00
Breusch-Pagan Test	BP = 5.67	0.13	4.53	0.21
Jarque-Bera Test	χ^2 = 71337	0.00	χ^2 = 69263	0.00
RESET Test	RESET = 0.43	0.63	RESET = 0.55	0.73

(Source: Own elaboration)

Figure 26 illustrates the autocorrelation function of the tracking errors for ESG and non-ESG ETFs. Even though the autoregressive pattern is visible in both cases, ESG ETFs appear to have weaker time-dependent relationships than non-ESG ETFs. Nevertheless, the ACF suggests that a dynamic model, such as GMM, might be more suitable for capturing these short-term relationships between current and historical tracking errors.

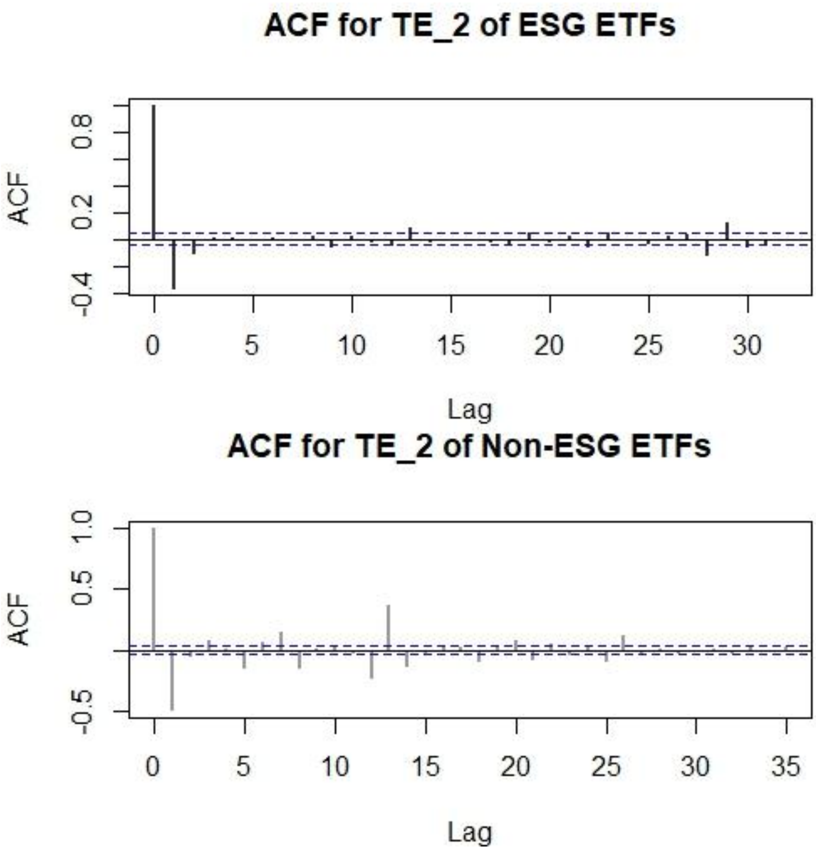

Figure 26. Autocorrelation Function (ACF) for TE_2 of ESG and Non-ESG ETFs
(Source: Own elaboration)

The VIF test results presented in Table 42 demonstrate that all model variables show no signs of multicollinearity. However, the DWH test indicates potential endogeneity issues. Endogeneity occurs when explanatory variables show correlation with the error term because of omitted variables, measurement errors, or reverse causality (Wooldridge, 2010). The model likely experiences endogeneity because it fails to account for autocorrelation. As a consequence of the violation of exogeneity assumptions, the coefficient estimates are biased and inconsistent.

Table 42. Tests for endogeneity and multicollinearity in the Random Effects models (ESG and non-ESG ETFs)

Explanatory Variable	ESG ETFs			Non-ESG ETFs		
	VIF	DWH Statistic	DWH p-value	VIF	DWH Statistic	DWH p-value
TER	-	24.17	0.00	-	24.12	0.00
AUM	1.00	26.22	0.00	1.00	26.29	0.00
RISK	1.00	24.97	0.00	1.00	24.91	0.00
AGE	1.00	24.18	0.00	1.00	24.15	0.00

(Source: Own elaboration)

To summarize, based on a preliminary analysis of the specificity of the data for ESG and non-ESG ETFs, it may be concluded that static models, including Random Effects, fail to account for the significant time dependence and endogeneity observed in the data. Accordingly, the tracking error for ESG and non-ESG ETFs is estimated using the GMM model structure that includes a first-order lagged dependent variable (TE_2,t-1), individual fixed effects, first-difference transformation, and one-step Arellano-Bond estimation with robust HAC standard errors. However, to reflect differences in data structure and autocorrelation patterns, the selection of instrumental variables was adjusted separately for ESG and non-ESG subsamples to ensure instrument validity.

To address autocorrelation in ESG ETFs, a set of instrumental variables was tested. As shown in Table 43, the exogeneity test (Hansen J test) confirms that all tested lags meet the exogeneity condition. However, the AR(2) test for lag (TE_2, 2) returned calculation errors, suggesting that this lag is too closely related to the current value of the dependent variable and may not be fully exogenous. Thus, to avoid the problem of overfitting, the GMM model uses lag (TE_2, 3) as an instrument. This variable is strongly correlated with TE_2 but remains uncorrelated with the error term, making it a valid and reliable instrument.

Table 43. Evaluation of lagged TE_2 as an instrumental variable in the GMM model (ESG ETFs)

Lag TE_2	Correlation with TE_2	AR(2) Test Statistic	AR(2) p-value	Hansen J-Test Statistic	Hansen p-value
2	0.35	–	–	47.99	1.00
3	0.16	-0.32	0.75	47.95	1.00
4	0.15	0.10	0.92	47.92	1.00
5	0.16	0.65	0.52	47.99	1.00
6	0.11	0.48	0.63	47.99	1.00
7	0.08	-0.80	0.42	47.95	1.00
8	0.07	-0.10	0.92	47.86	1.00
9	0.08	-1.03	0.30	47.89	1.00
10	0.03	1.05	0.29	47.99	1.00

(Source: Own elaboration

Table 44. Evaluation of lagged TE_2 as an instrumental variable in the GMM model (non-ESG ETFs)

Lag TE_2	Correlation with TE_2	AR(2) Test Statistic	AR(2) p-value	Hansen J-Test Statistic	Hansen p-value
2	0.34	0.26	0.84	116.25	0.01
3	0.37	-0.94	0.33	128.55	0.00
4	0.34	-0.13	0.87	124.15	0.00
5	0.35	0.80	0.43	123.76	0.01
6	0.33	0.24	0.81	120.65	0.01
7	0.33	-0.92	0.34	127.96	0.00
8	0.31	-0.13	0.87	124.15	0.00
9	0.28	0.80	0.43	123.76	0.01
10	0.37	0.20	0.85	126.08	0.00
11	0.39	-1.05	0.28	128.74	0.00
12	0.25	-0.12	0.88	123.86	0.01
13	0.28	-0.64	0.50	126.22	0.00
14	0.28	0.92	0.36	121.18	0.01
15	0.23	-0.76	0.43	122.81	0.01
16	0.23	-1.93	0.04	123.41	0.01
17	0.44	0.24	0.81	121.13	0.01
18	0.26	-2.86	-0.01	128.29	0.01
19	0.22	-1.44	0.13	117.89	0.03
20	0.22	-0.24	0.79	118.05	0.04
21	0.19	0.08	0.94	109.39	0.15
22	0.14	-0.10	0.89	113.31	0.08
23	0.19	-1.20	0.21	107.80	0.19
24	0.20	-0.15	0.86	107.54	0.26

(Source: Own elaboration)

Then, lag (TE_2, 24) was selected as the instrumental variable for the non-ESG ETF GMM model (see Table 44). While its correlation with the dependent variable is moderate (0.20), it presents the most favorable diagnostic profile: no evidence of second-order autocorrelation (AR (2) p = 0.95) and an acceptable Hansen J-test p-value (0.13), indicating instrument validity. In contrast, earlier lags, despite higher correlations, consistently failed the Hansen test.

The GMM estimation results for ESG ETFs, presented in Table 45, show that the key variables that affect TE_2 are its first lag, TER, AUM, and RISK. They confirm the dynamic nature of the tracking error process, with a statistically significant and positive coefficient of the lagged dependent variable. Then, the estimation results show that AUM has a negative and statistically significant effect on TE_2, indicating that larger ESG funds are associated with lower tracking error. In contrast, TER and RISK have positive and significant coefficients, suggesting that higher operating costs and benchmark volatility increase tracking error. The AGE variable is not statistically significant, which means that fund age does not have a clear effect on tracking error in ESG ETFs.

Table 45. Estimation results of the GMM model (ESG ETFs)

Model Formula	TE_2 ~ lag (TE_2, 1) + TER + AUM + RISK + AGE | lag(TE_2, 3)			
Variable	**Estimate**	**Std error**	**Statistic**	**p-value**
lag(TE_2, 1)	0.16	0.05	3.03	0.00
TER	0.02	0.01	1.81	0.07
AUM	-0.03	0.01	-3.48	0.00
RISK	0.05	0.01	3.47	0.00
AGE	0.03	0.02	1.51	0.13

(Source: Own elaboration)

In contrast, the GMM estimation results for non-ESG ETFs presented in Table 46 reveal a slightly different pattern.

Table 46. Estimation results of the GMM model (non-ESG ETFs)

Model Formula	TE_2 ~ lag (TE_2, 1) + TER + AUM + RISK + AGE | lag(TE_2, 3)			
Variable	**Estimate**	**Std error**	**Statistic**	**p-value**
lag(TE_2, 1)	0.10	0.05	1.93	0.05
TER	0.06	0.02	3.10	0.00
AUM	0.08	0.05	1.60	0.11
RISK	0.03	0.01	2.74	0.01
AGE	-0.02	0.02	-1.19	0.24

(Source: Own elaboration)

In the case of non-ESG ETFs, tracking error is influenced by its past values, although the effect of the lagged variable is weaker than in the ESG group. TER has a positive and significant impact, meaning that higher fund costs reduce tracking precision. Unlike in the ESG sample, AUM is not statistically significant, which suggests that fund size does not play an important role in tracking performance here. RISK is positively and significantly related to TE_2, while AGE has no significant effect.

The diagnostic tests in Table 47 validate the GMM model. The AR(1) test results confirm the expected first-order autocorrelation, and the AR(2) test indicates no second-order autocorrelation. The Hansen test shows that the instruments are correctly specified. According to the results of the Wald test, all explanatory variables are jointly significant. These results indicate that the model is correctly specified, and the estimates are reliable.

Table 47. Diagnostic tests of the GMM model (ESG and non-ESG ETFs)

Test	**ESG ETFs**		**Non-ESG ETFs**	
	Statistic	**p-value**	**Statistic**	**p-value**
Autocorrelation Test AR (1)	-3.67	0.00	-4.67	0.00
Autocorrelation Test AR (2)	-0.32	0.75	-0.64	0.52
Hansen Test	$\chi^2 = 47.95$	0.99	$\chi^2 = 116.97$	0.22
Wald Test	$\chi^2 = 126.74$	0.00	$\chi^2 = 86.45$	0.00

(Source: Own elaboration)

Based on the results, the specific hypothesis HS1 is confirmed. The lagged dependent variable is statistically significant in both models, confirming the autoregressive nature of tracking errors for both ESG and non-ESG ETFs. HS2 is also supported, as TER shows a positive effect on tracking error in both groups. HS3 is not confirmed. Even though AUM significantly reduces tracking error in ESG ETFs, the relationship is not statistically significant in the non-ESG group. HS4 is not supported since fund age does not show a significant impact on tracking error in either model. Finally, HS5 is confirmed, as benchmark volatility is positively associated with tracking error in both ESG and non-ESG ETFs.

Table 48 provides the results of the Chow test. It is used to verify whether the relationships between variables in the model are the same for ESG and non-ESG ETFs. The analysis reveals that the two groups have partial structural differences. The Chow test indicates that AUM demonstrates a significant difference between ESG and non-ESG ETFs, which means that fund size impacts tracking error differently in these two groups. Next, there is a weak but significant difference (10% significance level) for TER, which suggests some variation in cost effects on the tracking errors of ESG and non-ESG ETFs. The analysis reveals no meaningful differences between RISK and the lagged dependent variable, while the estimate of the AGE variable was not significant. Overall, the findings from this study partially validate hypothesis H2, suggesting that the determinants of tracking error might differ between passive ESG equity ETFs listed on European exchanges and their non-ESG counterparts.

Table 48. Chow test results for GMM models (ESG vs. non-ESG ETFs)

Variable	ESG Estimate	Non-ESG Estimate	Z Statistic	p-value
lag (TE, 1)	0.16	0.10	0.85	0.40
TER	0.02	0.06	-1.79	0.07
AUM	-0.03	0.08	-2.16	0.03
RISK	0.05	0.03	1.41	0.16
AGE	0.03	-0.02	-	-

(Source: Own preparation)

The estimation results presented in Table 49 verify the ESG factor's moderating influence on ETF tracking error determinants. The model introduces the interaction terms with the ESG factor for only those variables that showed significant differences in the Chow test. The unified model allows for direct statistical comparison between groups, which provides a more precise and controlled assessment of potential group-level differences. The results of standard diagnostic tests appear in Table 50, which validates the model and strengthens the stability of the research findings.

Table 49. GMM model estimation with ESG interaction terms (full ETF sample)

Model Formula	TE_2 ~ lag (TE_1, 1) + TER + AUM + RISK + AGE + ESG + REP + ESG: TER + ESG: AUM +| lag (TE_1,20)			
Variable	**Estimate**	**Std error**	**Statistic**	**p-value**
ESG	-0.02	0.01	-3.50	0.00
lag (TE, 1)	0.19	0.10	1.85	0.06
TER	0.02	0.01	1.47	0.14
AUM	0.03	0.03	0.79	0.43
RISK	0.00	0.01	0.22	0.83
AGE	0.01	0.02	0.66	0.51
REP	-0.02	0.01	-3.52	0.00
ESG: TER	-0.03	0.03	-0.89	0.37
ESG: AUM	-0.08	0.06	-1.38	0.17

(Source: Own elaboration)

The negative and statistically significant coefficient on the ESG variable shows that, on average, ESG ETFs have lower tracking errors than non-ESG ETFs, after accounting for other factors. The positive and significant coefficient on the lagged tracking error means that tracking error is persistent over time. Finally, the negative and significant effect of the synthetic replication variable (REP) suggests that this replication method improves tracking accuracy.

However, in this model estimated for the full sample, variables such as TER, AUM, and RISK are no longer statistically significant. This is probably due to higher variance in the combined sample and the presence of interaction terms (Greene, 2008). Importantly, none of the interaction terms between ESG and the key explanatory variables are statistically significant. This means that the effects of TER and AUM on tracking error do not differ meaningfully between ESG and non-ESG ETFs.

Table 50. Diagnostic tests of the model with ESG interaction terms (full ETF sample)

Test	**Statistic**	**p-value**
Autocorrelation Test AR(1)	-3.57	0.00
Autocorrelation Test AR(2)	-0.44	0.66
Hansen Test	χ^2 = 112.74	0.22
Wald Test	χ^2 =314.11	0.00

(Source: Own elaboration)

In conclusion, according to the Chow test results, there are potential differences in tracking error determinants between ESG and non-ESG ETFs, which suggests these funds operate through slightly different mechanisms. On the other hand, the interaction model fails to confirm this preliminary indication, as it provides a more direct and statistically robust test of such differences. The lack of consistency between the two approaches means there is no strong empirical basis to conclude that the determinants of tracking error differ across groups. The observed differences should be viewed with caution because they remain weak and statistically unreliable.

4. 5 Discussion

The results of this study show that passive equity ETFs listed on European exchanges achieve high tracking accuracy, regardless of whether they follow ESG indices or not. Out of 134 analysed ETFs, only four slightly exceeded the 0.5% tracking error threshold commonly used in the literature (Banerjee, 2015), and none went above 0.53%. This corresponds with previous research that ETFs exhibit low tracking errors (Frino and Gallagher, 2001; Elton et al., 2002; Agapova, 2011; Rompotis, 2011; Chu, 2011).

This study adds new evidence on ESG ETFs. The results confirm that ESG ETFs replicate their benchmarks effectively, in line with the results of Nguyen (2023), and that their tracking error remains stable over time (Lee, 2020). These findings suggest that the ESG factor does not weaken the ability of ETFs to closely follow their indices.

While the cross-sectional analysis found no significant difference in tracking error between ESG and non-ESG ETFs, the dynamic panel model (GMM) indicated otherwise. In that model, the ESG dummy was statistically significant, pointing to a systematic difference in tracking performance. This difference in results may be explained by the fact that cross-sectional models do not fully capture time-based patterns, whereas panel models account for dynamic effects more effectively (Bowen and Wiersema, 1999).

The first research hypothesis (H1), which states that there is a significant difference between the tracking errors of passive broad market ESG equity ETFs listed on European exchanges and their non-ESG counterparts, is supported by the results. ESG ETFs show slightly lower tracking errors on average, indicating better benchmark replication. This suggests that ESG strategies do not weaken ETF performance and may even offer an advantage. One possible reason for this result is the strong competition in the ETF market, which may push ESG ETF managers—operating in a fast-growing segment—to deliver high tracking accuracy to meet investor expectations (Box et al., 2018).

This study delivers new evidence on what drives tracking error in passive equity ETFs listed on European exchanges. The strongest effect was observed for lagged tracking error, which supports hypothesis HS1. This confirms that tracking error follows an autoregressive pattern in both ESG and non-ESG ETFs. This result is in line with Ivanov (2015), who found a similar pattern in currency ETFs using autoregressive models. The results show that when fund managers find effective ways to track an index, they continue to use them. This consistent approach helps keep tracking performance stable and reliable over time.

The research evidence supports hypothesis HS2, which demonstrates that higher total expense ratios (TER) result in greater tracking errors. The relationship is consistent with the results of Frino and Gallagher (2001) and Blitz et al. (2012). Therefore, the study confirms that maintaining low costs continues to be vital for achieving precise index replication.

Regarding fund size, the results for ESG and non-ESG funds were inconsistent. In ESG ETFs, larger fund size (AUM) was linked to lower tracking error. In non-ESG ETFs, this effect was not statistically significant. Because the expected relationship was not found in non-ESG ETFs, hypothesis HS3 was rejected. These results differ from Chu (2011), who found that in Hong Kong, ETFs with more assets tracked their index more accurately, no matter what type of fund they were. The size of assets non-ESG ETFs in Europe may already be large, and additional growth does not lead to improved tracking efficiency. Then, ESG ETFs could still achieve better results through increased scale. This means that the tracking accuracy of ESG fund managers would improve when they expand their fund size.

The study found no significant relationship between ETF age and tracking error. As a result, hypothesis HS4, which proposed that there is a negative relationship between fund age and the tracking error of both ESG and non-ESG passive equity ETFs listed on European exchanges, was rejected. This finding contrasts with earlier studies by Rompotis (2011) and Chu (2011), which emphasized the role of fund age in tracking performance. The results show that fund age no longer plays a major role in explaining ETF tracking quality. This may be due to the growing standardization and professionalism of ETF management in Europe (ESRB, 2019). ESMA guidelines have improved transparency, reporting, and portfolio management rules (Lansing, 2013). As a result, both older and newer ETFs now operate more similarly, making fund age less important for tracking performance.

Finally, the results support hypothesis HS5, which suggests that there is a positive relationship between benchmark volatility and the tracking error of both ESG and non-ESG passive equity ETFs listed on European exchanges. This finding is in line with earlier studies showing that ETFs tend to have higher tracking errors during periods of greater market volatility (Rompotis, 2011; Qadan and Yagil, 2012; Drenovak et al., 2014; Vardharaj et al., 2004).

The results do not provide sufficient evidence to confirm hypothesis H2, which stated that the determinants of tracking error differ between passive ESG equity ETFs listed on European exchanges and their non-ESG counterparts. The Chow test showed some possible differences, especially in the role of fund size and costs. However, these differences were not confirmed in the interaction model, which gives a more direct and reliable test. The interaction terms were not significant, meaning that the effects of key factors like TER and AUM are similar for both ESG and non-ESG ETFs. Therefore, any differences found should be seen as weak and not statistically reliable.

Even though the study did not show differences in tracking error determinants between ESG and non-ESG ETFs, the significant effect of the ESG factor was confirmed in all models. The negative coefficient of the ESG variable may suggest that it is easier for fund managers to track ESG indices rather than the traditional ones. In that way, the study challenges the assumption that ESG and non-ESG indices are functionally equivalent regarding replicability. It extends classical passive investment theory to the context of sustainable finance.

The results of this study complement previous research on the different mechanisms of ESG vs. traditional indices. Empirical evidence on the characteristics of ESG indices shows significant differences in comparison to traditional indices. For instance, Avramov et al. (2022) argue that ESG integration stabilizes cash flows and reduces risk, whereas Pedersen et al. (2021) point out potential diversification losses resulting from ESG-related exclusions. Plastun et al.

(2025) documented that ESG indices require distinct forecasting models from traditional indices, potentially due to higher information transparency and lower speculative activity.

The improved tracking performance of ESG ETFs relative to traditional ETFs may stem from the construction methods of ESG indices. The exclusion of companies with high ESG risks stands as a key feature of ESG indices because they remove companies from controversial industries and those with poor governance practices. These exclusions significantly reduce the number of volatile stocks in the index, which results in lower company-specific risk (Choueifaty et al., 2023). Finally, reduced idiosyncratic risk leads to more stable returns, which facilitates index replication by ETF managers (MSCI, 2022). The ESG indices use sustainability scores to determine which companies receive more weight in an index. The selection process tends to favor companies with stable financial performance and lower market volatility (Bennani et al., 2018). In that way, the structural characteristics of ESG ETFs help explain their enhanced tracking performance and replication consistency compared to non-ESG ETFs.

The study results should be interpreted with some important limitations in mind. The research focuses on European exchange-traded funds. This is justified because of the European leadership in ESG investing and regulation. However, it also means that the results may not translate to US or Asian markets because their ESG rules and market conditions differ. What is more, the study period spans less than four years. In that way, the research captures the recent development of ESG ETFs and examines ESG investing changes from the COVID-19 pandemic period until the present day. However, due to the selected timeframe, the study fails to demonstrate how performance tracking evolves across extended periods.

Finally, the study includes only simple equity ETFs. By including in the sample only plain-vanilla ETFs, the study focuses on standard, uncomplicated funds that aim to track the performance of a broad stock market index, without using complex strategies. More sophisticated funds, such as smart-beta or leveraged ETFs, were excluded to keep the analysis clear and consistent. For these reasons, the conclusions mainly refer to plain-vanilla passive equity ETFs in the European market and should not be seen as fully general.

Future studies could examine the tracking performance of ESG ETFs in other geographic regions. This would allow the verification of whether the results observed in Europe appear in different markets. Then, researchers could examine ESG ETFs over a longer period to see if tracking errors stay the same or change when market conditions shift. Studying different types of ESG strategies—such as exclusion-based, thematic, or multi-factor—would help explain how the way an index is built affects ETF tracking. It would also be useful to analyse ESG ETFs over full market cycles, from crisis to recovery, to determine how well they perform under stress and whether they can still follow their benchmark accurately.

4. 6 Summary

This chapter of the dissertation is focused on the empirical examination of the tracking ability of 48 passive broad market ESG equity ETFs listed on European exchanges compared to their 86 non-ESG counterparts over the period from January 1, 2021, to June 30, 2024. Moreover, it explored the key determinants of tracking errors of ESG and non-ESG ETFs.

The research shows that ESG and non-ESG ETFs both exhibit excellent tracking ability. The cross-sectional analysis revealed no statistically significant differences in tracking error between these two groups. Then, the dynamic panel model, which introduced the time component, showed that the ESG variable had a significant impact on the tracking error of ETFs. The findings confirm hypothesis H1, which proposed that passive ESG equity ETFs listed on European exchanges exhibit significantly different tracking errors compared to their non-ESG counterparts. The results show a weak but consistent improvement in replication accuracy for ESG ETFs.

The GMM model revealed essential factors that affect the tracking error of ETFs. The results showed that past tracking errors strongly determine their future tracking errors, revealing their autoregressive nature. The results indicated that benchmark volatility and the fund’s total expense ratio (TER) affected tracking error in both ESG and non-ESG ETFs. Then, assets under management significantly affect tracking errors only in ESG ETFs, while fund age has no significant impact on either group.

The Chow test confirmed structural differences in how fund size and cost impact the tracking error across the two groups. However, the model that included interaction terms failed to confirm stable differences between groups. Therefore, the hypothesis H2, which stated that the determinants of tracking error differ between passive ESG equity ETFs listed on European exchanges and their non-ESG counterparts, was rejected. The core determinants of tracking error appear to be similar between ESG and non-ESG ETFs.

CONCLUSION

ESG and passive investing are two revolutionary trends in the international financial market. These two trends are combined in ESG exchange-traded funds (ETFs), which track the performance of indices built in line with the sustainability criteria. The rising importance of ESG investing in Europe shows that investors are shifting their attention to sustainability and responsible investment practices. They are becoming more aware of global environmental and social challenges, as well as new regulations that promote responsible and ethical finance. Frameworks such as the Sustainable Finance Disclosure Regulation (SFDR) and the EU Taxonomy have accelerated the design of ESG financial products. ESG ETFs have become important tools for sustainable finance, offering investors a way to align financial goals with broader social and environmental values. This dissertation contributes to this evolving landscape by investigating the relationship between passive portfolio management and ESG investing, with a particular focus on the tracking ability of passive equity ETFs listed on European exchanges.

The main objective of this dissertation included the evaluation and comparison of the tracking ability of passive ESG equity ETFs listed on European exchanges, in relation to their non-ESG counterparts, as well as the identification of the key factors affecting replication quality. The study examined 134 ETFs, including 48 ESG and 86 non-ESG ETFs, using weekly data from January 2021 to June 2024. The research employed dynamic panel regression models (Arellano–Bond GMM) to handle both autocorrelation and endogeneity issues.

The research results revealed that passive ESG equity ETFs listed on European exchanges exhibit significantly different tracking errors compared to their non-ESG counterparts. The GMM model results showed that the integration of ESG criteria changes the replication performance of ETFs. The negative coefficient of the ESG variable indicates that ESG ETFs tend to have lower tracking errors, meaning they replicate their benchmark indices more precisely than non-ESG funds.

The dissertation confirms the applicability of traditional tracking ability determinants to ESG ETFs and indicates some nuances that differentiate them from non-ESG ETFs. Firstly, the study confirms the autoregressive character of ETF tracking errors and the relevance of total expense ratio (TER) and benchmark risk as universal factors shaping tracking errors of both ESG and non-ESG ETFs. The research also showed some distinct patterns of how structural elements impact ESG and non-ESG ETF tracking precision. The analysis confirmed that ESG funds benefit from larger asset bases, which enhance their replication precision. The non-ESG funds did not show this effect because their tracking performance does not seem to be influenced by the fund size. The difference between ESG and non-ESG ETFs shows that ESG funds gain more from growing their asset size. As a newer and developing group, they likely use these additional assets to improve liquidity, lower costs, and track the index more precisely. In contrast, most non-ESG ETFs are already large and well-established, so having more assets does not noticeably improve their replication performance.

Hypothesis H2 assumed that the factors influencing tracking error differ between passive ESG and non-ESG equity ETFs listed in Europe. The results do not provide clear support for this claim. The Chow test indicated some variation. In particular, costs (TER) had a stronger impact on non-ESG funds, while fund size (AUM) helped reduce tracking error only in ESG funds. However, the initial findings from the Chow test were not confirmed by the more comprehensive interaction model. The model tested the interaction terms with the ESG factor for only those variables that showed significant differences in the Chow test. It showed that all interaction effects proved statistically insignificant. This means that TER and AUM did not produce different effects on the tracking error of ESG vs. non-ESG ETFs.

Although the ESG variable itself remained significant and negative, indicating that ESG ETFs tend to have lower tracking errors on average, this reflects a general difference in performance level, not a difference in the way structural factors influence that performance. Therefore, despite the Chow test suggesting some group-level variation, the interaction model offers no strong empirical evidence that ESG and non-ESG ETFs are shaped by different mechanisms when it comes to the drivers of tracking error. As a result, Hypothesis H2 was rejected, and any observed differences should be interpreted with caution, as they are not robust across methodological approaches.

From a theoretical perspective, this dissertation adds to the growing research on passive ESG investing. The results show that environmental, social, and governance (ESG) factors affect how ETFs track their benchmark indices. This may suggest that ESG indices are easier for fund managers to track, likely due to the distinct construction methods of ESG indices, as compared to the non-ESG ones. As noted by Choueifaty et al. (2023), the removal of companies with high ESG risk may lead to a decrease in index volatility, which may suggest that such companies are characterized by higher individual volatility. Because of that, ESG criteria can affect how well passive funds follow the underlying benchmark index. This finding helps to explain the role of ESG in managing passive investment funds.

From a practical point of view, the research provides useful information to both investors and ETF providers. The tracking performance of ESG ETFs makes them an appropriate choice for investors who pursue a sustainable and cost-effective alternative to non-ESG funds. The dissertation also challenges the results of some studies (Rompotis, 2011; Chu, 2011) about fund age importance while confirming the importance of universal tracking error determinants. The factors that determine replication quality consist of historical tracking performance, costs, together with benchmark volatility, replication method, and the ESG status of the underlying benchmark.

The research suggests that ESG ETFs managers need to expand their fund size to achieve improved replication accuracy. The achievement of this goal demands persistent investor engagement through transparent communication about performance metrics, structures, and portfolio composition. The strategic growth of ESG ETFs requires specific promotional efforts to reach the target audience. Promoting passive ESG investing means clearly explaining its main strengths, such as low fees, alignment with sustainable goals, and the ability to follow index performance with accuracy. The ESG ETF market can grow faster if these funds are easier to access, for example, by listing them on more stock exchanges or offering them on global investment platforms.

The wider use of ESG ETFs is still difficult because ESG standards are not clearly defined or applied in the same way across the market. When rules and strategies are vague, investors may feel uncertain, which lowers their trust in these funds. To improve transparency, index providers should clearly explain how they build ESG indices, how they assess companies, and what rules they follow when deciding which firms to include or exclude. ETF managers should collaborate with index providers who apply these transparent practices. Building long-term trust in ESG investing depends on clarity, openness, and consistency.

This dissertation offers investors clear and practical guidance. The findings show that ESG ETFs replicate their benchmark indices with a high level of accuracy. As a result, investors can confidently include ESG ETFs in their portfolios without compromising on tracking quality. Investors should apply the same evaluation process to ESG ETFs that they would use for traditional passive funds. The assessment of ESG ETFs requires evaluation of their historical tracking error performance, together with their total expense ratio and benchmark index risk level. It is also important to consider the size of the fund. ESG ETFs with larger assets tend to have lower tracking errors, which means that they follow their index more accurately.

This dissertation also offers useful conclusions for policymakers. They have a key role in improving transparency and reporting standards for ESG ETFs. Clear and consistent information is essential to build investor trust and support the further growth of ESG funds. Stronger regulation, especially in the area of mandatory disclosures, can help reduce uncertainty. This includes clear reporting on how ESG strategies are applied, what data and methods are used, and what companies are included in the portfolio. Better rules in these areas would help investors understand what stands behind the ESG label and how it may affect fund performance. In addition, more consistent rules across Europe could make the market more accessible and competitive by setting common standards for index providers and ETF issuers.

The research contains certain limitations in its design. The research focuses on passive equity ETFs operating in Europe during the period from January 2021 through June 2024. The analysis excludes more sophisticated fund types, including leveraged, inverse, and currency-hedged ETFs. The results obtained from this study do not represent the entire ETF market. Therefore, divergent outcomes might emerge when studying other fund types or regions. However, this narrow focus was chosen on purpose. It allowed for a reliable and clear comparison between ESG and non-ESG funds and enabled the verification of how the ESG factor affects tracking performance. The period matches the early stage of the new ESG regulation in the EU (SFDR), which reduced the risk of greenwashing and made ESG classifications more reliable. Although the scope is limited, it makes the comparison more accurate.

These limitations point to some areas worth exploring in future research. A longer study period could show whether tracking performance changes when market conditions shift. Including other types of ETFs, such as sector, thematic, bond, or commodity funds, may help explain how ESG affects different index structures. Research on ETFs from other regions, like North America or Asia, could show how local regulations influence replication. Adding more factors, such as fund liquidity, trading costs, or economic conditions, could lead to a better understanding of what causes tracking error.

The growing importance of ESG integration within passive investing frameworks reflects a broader shift in how capital markets respond to environmental, social, and governance challenges. This dissertation improves the understanding of passive ESG investing and supports the further development of the ESG ETF market. The results show that investors can invest in passive ESG equity ETFs listed in European exchanges without compromising the quality of index replication. Concerning the tracking ability determinants, the research confirmed that the factors influencing the tracking error of ESG and non-ESG ETFs are largely similar. Given the growing importance of ESG regulation and investor interest, these findings are relevant for fund managers, regulators, and market participants. Future research should continue to explore how ESG integration affects the functioning of passive funds to support transparent, efficient, and scalable investment solutions.

BIBLIOGRAPHY

Abner, D. (2016). The ETF Handbook: How to Value and Trade Exchange-Traded Funds. Hoboken, N.J. : Wiley.

Ackert, L. F. and Tian, Y. S. (2000). Evidence on the efficiency of index options markets. Economic Review, Federal Reserve Bank of Atlanta, 85(Q1), 40-51.

Ackert, L. F. and Tian, Y. S. (2008). Arbitrage, Liquidity, and the Valuation of Exchange Traded Funds. Financial Markets, Institutions & Instruments, 17, 331-362.

Adamsson, H. and Hoepner, A. G. (2015). The Price of Sin' Aversion: Ivory Tower Illusion or Real Investable Alpha? SSRN Electronic Journal Available at: https://ssrn.com/abstract=2659098.

Agapova, A. (2011). Conventional mutual index funds versus exchange-traded funds. Journal of Financial Markets,14 (2), 323-343.

Alford A. and Guffey D. M. (1996). A Re-examination of International Seasonalities, Review of Financial Economics, 5, 1-17.

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulats et axioms de l'école américaine. Econometrica, 21, 503–46.

Amenc, N. and Martellini, L. (2003). The Brave New World of Hedge Fund Indices. Working Paper, Risk and asset management research centre, Edhec Business School, Nice.

An, Y., Benetton, M. and Song, Y. (2021). Index Providers: Whales Behind the Scenes of ETFs. St. John's University School of Law Legal Studies Research Paper Series.

Andreu, J. (2009). Market Indices: Bases, Biases and Beyond. PhD Thesis.

Antolín-López, R. and Ortiz-De-Mandojana, N. (2023). Measuring and disclosing Environmental, Social and Governance (ESG) information and performance, EUR 31372 EN, Publications Office of the European Union, Luxembourg.

Antoniewicz, R. and Heinrichs, J. (2014). Understanding Exchange-Traded Funds: How ETFs Work. SSRN Electronic Journal. Available at: https://ssrn.com/abstract=2523540.

Arnott, R., Hsu, J. and Moore, P. (2005). Fundamental Indexation. Financial Analysts Journal, 61 (2), 83-99.

Arya, S. and Kaplan, P. (2006). Collared Weighting: A Hybrid Approach. Morningstar Indexes Yearbook, 3, 22–25.

Asness, C. S., Moskowitz, T. J. and Pedersen, L. (2013). Value and Momentum Everywhere. Journal of Finance, 68(3), 929-985.

Auer, B. R. (2016). Do Socially Responsible Investment Policies Add or Destroy European Stock Portfolio Value? Journal of Business Ethics, 135(2), 381–397.

Bachelier, L. (1900). Theorie de la speculation. Annales scientifiques de l'école normale supérieure, 17, 21–86.

Baker, A., Larcker, D. and Wang, C. (2022). How much should we trust staggered difference-in-differences estimates? Journal of Financial Economics, 144(2), 370-395.

Baker, H. K. and Nofsinger, J. R. (2010). Behavioral Finance: Investors, Corporations, and Markets. Hoboken, New Jersey: John Wiley & Sons, Inc.

Ball, R. and Brown, P. (1968) An Empirical Evaluation of Accounting Income Numbers. Journal of Accounting Research, 6, 159-178.

Bansal, R., Wu, D. and Yaron, A. (2022). Socially Responsible Investing in Good and Bad Times. Review of Financial Studies, 35(4), 2067-2099.

Barnea, A. and Rubin, A. (2010). Corporate Social Responsibility as a Conflict Between Shareholders. Journal of Business Ethics, 97(1), 71-86.

Barras, L., Scaillet, O., and Wermers, R. (2010). False Discoveries in Mutual Fund Performance: Measuring Luck in Estimated Alphas. The Journal of Finance, 65(1), 179–216.

Battalio, R. H. and Mendenhall, R. R. (2005). Earnings expectations, investor trade size, and anomalous returns around earnings announcements. Journal of Financial Economics, Elsevier, 77(2), 289-319.

Baum, C. F., Schaffer, M. E., & Stillman, S. (2003). Instrumental Variables and GMM: Estimation and Testing. Stata Journal, 3(1), 1–31.

Bebchuk, L. and Hirst, S. (2022). Big Three Power, and Why it Matters. Boston University Law Review, 102, 1547-1600.

Bebchuk, L., Cohen, A. and Wang, C. (2013). Learning and the Disappearing Association Between Governance and Returns. Journal of Financial Economics, 108 (2), 323-348.

Becher, D. and Campbell, T. (2004). Corporate governance of bank mergers. Federal Reserve Bank of Chicago Proceedings, 918.

Beine, M., Cosma, A. and Vermeulen, R. (2010). The dark side of global integration: Increasing tail dependence. Journal of Banking & Finance, 34(1), 184-192.

Bekaert, G., Hodrick, R. J. and Zhang, X. (2009). International Stock Return Comovements. The Journal of Finance, 64(6), 2591–2626.

Bennani, L., Le Guenedal, T., Lepetit, F., Ly, L., Mortier, V., & Sekine, T. (2018). How ESG investing has impacted the asset pricing in the equity market. Amundi Asset Management.

Berg, F., Kölbel, J. F Rigobon, R. (2022). Aggregate Confusion: The Divergence of ESG Ratings. Review of Finance, 26 (6), 1315–1344.

Berger, J. M. and Mandelbrot, B. (1963). A new model for error clustering in telephone circuits. IBM Journal of Research and Development ,7(3), 224–236.

Bhattacharya, U. and Galpin, N. (2011). The Global Rise of the Value-Weighted Portfolio. The Journal of Financial and Quantitative Analysis, 46(3), 737–756.

BlackRock (2010). Overview of ETF Structures. How ETFs track their benchmark.

BlackRock (2022). Enhanced Disclosures by Certain Investment Advisers and Investment Companies About Environment, Social, and Governance Investment Practices", Securities Act No. 11068; Securities Exchange Act No. 94985; Investment Company Act No.34594; and Advisers Act No. 6034, 87 Fed. Reg. 117 (June 17, 2022), 36654 ("Proposing Release").

Blake, D. and Timmermann, A. (1998). Mutual Fund Performance: Evidence from the UK. Review of Finance, European Finance Association, 2(1), 57-77.

Blitz, D. and Huij, J. (2012). Evaluating the performance of global emerging markets equity exchange-traded funds. Emerging Markets Review, 13(2), 149-158.

Blitz, D., Huij, V. and Swinkels, L. (2012). The Performance of European Index Funds and Exchange Traded Funds. European Financial Management, 18(4), 649-662.

Bloomberg (2022), ESG Stripped From 23% of EU's Sustainable Funds in Fresh Review. Retrieved from: https://www.bloomberg.com/news/articles/2022-08-18/-esg-stripped-from-23-of-eu-sustainable-funds-in-fresh-review.

Boddy, R. and Smith, G. (2009). Statistical Methods in Practice: For scientists and technologists. Chichester, U.K.: Wiley.

Bodie, Z., Kane, A., & Marcus, A. J. (2013). Essentials of Investments (9th ed.). McGraw-Hill/Irwin.

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

Bogle, J. C. (2007). The Little Book of Common Sense Investing: The Only Way to Guarantee Your Fair Share of Stock Market Returns. John Wiley & Sons, Inc.

Bolton, P. and Kacperczyk, M. (2021). Do investors care about carbon risk? Journal of Financial Economics, 142 (2), 517-549.
Braun, B. (2015). From performativity to political economy: index investing, ETFs and asset manager capitalism. New Political Economy, 21 (3), 257–273.
Brealey, R.A., Myers, S.C. and Allen, F. (2011). Principles of Corporate Finance. 10th Edition, New York: McGraw-Hill/Irwin.
Brentani, C. (2004). Portfolio Management in Practice. Oxford: Elsevier Butterworth-Heinemann.
Brown, S., Goetzmann, W.N., Ibbotson R. and Ross, S. (1992) Survivorship Bias in Performance Studies. Review of Financial Studies, 5, 553-580.
Brundtland, G. H. (1987). Our Common Future: Report of the World Commission on Environment and Development. Geneva, UN-Dokument A/42/427.
Bubb, R, and Catan, E. (2022). The Party Structure of Mutual Funds. Review of Financial Studies, 35(6), 2839-2878.
Buetow, G. W. and Henderson, B. J. (2012). An Empirical Analysis of Exchange-Traded Funds. Journal of Portfolio Management, 38, 112-127.
Bugg-Levine, A. and Emerson, J. (2011), Impact Investing: Transforming How We Make Money while Making a Difference. Innovations: Technology, Governance, Globalization, 6, (3), 9-18.
Buiter, W. (2003). James Tobin: An Appreciation of his Contribution to Economics. NBER Working Papers 9753, National Bureau of Economic Research, Inc.
Burton Tailor International Consulting (2022). Index Industry Global Share & Segment Sizing 2022.
Business and Human Rights Resource Centre (2019). UN Guiding Principles. Available at: https://www.business-humanrights.org/en/big-issues/un-guiding-principles-on-business-human-rights/.
Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. The Journal of Finance, 52, 57-82.
Carlson, R. S. (1970). Aggregate Performance of Mutual Funds, 1948-1967. Journal of Financial and Quantitative Analysis, 5, 1-32.
CFA Institute (2019). CFA Program Curriculum 2020 Level I Volume 4. Wiley.
CFA Society of the UK (2021). Certificate in ESG Investing Curriculum.
Chan L. and Felmingham, B. (2003). The interdependence of share markets in the developed economies of East Asia. Pacific-Basin Finance Journal, 11(2), 219-237.
Chang, E. C. and Lewellen, W. G. (1984) Market Timing and Mutual Fund Investment Performance. Journal of Business, 57, 57-72.
Charupat, N. and Miu, P. (2013). The Pricing Efficiency Of Leveraged Exchange-Traded Funds: Evidence From The U.S. Markets. Journal of Financial Research, 36(2), 253-278.
Chatterji, A. K., Durand, R., Levine, D. and Touboul, S. (2016). Do ratings of firms converge? Implications for managers, investors and strategy researchers. Strategic Management Journal, 37 (8), 1597-1614.
Chatterji, A. K., Levine, D. I., and Toffel, M. W. (2009). How well do social ratings actually measure corporate social responsibility? Journal of Economics and Management Strategy, 18(1), 125–169.
Chatterji, A., Durand, R., Levine, D. and Touboul, S. (2019). Do ratings of firms converge? Implications for managers, investors and strategy researchers. Strategic Management Journal, 37 (8), 1597-1614.
Cheema-Fox, A., LaPerla, B. R., Serafeim, G., Turkington, D. and Wang, H. (2021). Decarbonization Factors. The Journal of Impact and ESG Investing, 2 (1) 47-73.

Chen, J., Jiang, W. and Xiaolan, M. Z. (2022). The Economics of Mutual Fund Marketing. SSRN Electronic Journal, Available at: https://ssrn.com/abstract=4277000.

Choueifaty, Y., Froidure, T.,and Reynier, J. (2023). ESG and idiosyncratic volatility. Journal of Investment Strategies. https://link.springer.com/article/10.1007/s41471-023-00155-1

Chowdhury, D. (2004). Incentives, Control and Development: Governance in Private and Public Sector with Special Reference to Bangladesh. Dhaka: Viswavidyalav Prakashana Samstha.

Chowdhury, M., Howe, J. S. and Lin, J. (1993). The Relation between Aggregate Insider Transactions and Stock Market Returns. Journal of Financial and Quantitative Analysis, Cambridge University Press, 28(3), 431-437.

Christensen, M. V. (2005). Danish Mutual Fund Performance - Selectivity, Market Timing and Persistence. Finance Research Group Working Paper, F-2005-1.

Christoffersen, P., Errunza, V., Jacobs, K. and Langlois, H. (2012). Is the Potential for International Diversification Disappearing? A Dynamic Copula Approach. Review of Financial Studies, 25 (12), 3711-3751.

Chu, P. (2011). Study on the tracking errors and their determinants: evidence from Hong Kong exchange traded funds. Applied Financial Economics, 21(5), 309-315.

CICERO Shades of Green (2021). Best Practices.

Ciciretti, R., Dalò, A. and Dam, L. (2017). The Price of Taste for Socially Responsible

Ciciretti, R., Dalò, A., and Dam, L. 2023. The Contributions of Betas Versus Characteristics to the Esg Premium. Journal of Empirical Finance. 71 pp. 104-124.

Coeurdacier, N. and Guibaud, S. (2011). International portfolio diversification is better than you think. Journal of International Money and Finance, Elsevier, 30(2), 289-308.

Cootner, P. H. (ed.) (1964). The Random Character of Stock Market Prices. Cambridge: The MIT Press.

Cornell, B. and Damodaran, A. (2020). Valuing ESG: Doing Good or Sounding Good? The Journal of Impact and ESG Investing, 1, 76-93.

Coval, J. D. and Shumway, T. (2005). Do Behavioral Biases Affect Prices? Journal of Finance, 60(1), 1-34.

Cowles, 3rd, A. and Jones, H. E. (1937), Some a posteriori probabilities in stock market action,. Econometrica, 5(3), 280–294.

Cremers, M., Ferreira, M., Matos, P., and Starks, L. (2016). Indexing and active fund management: International evidence. Journal of Financial Economics, 120(3), 539–560.

Dalkir, M. (2009). Revisiting stock market index correlations. Finance Research Letters, 6(1), 23-33.

De Bondt W. F. and Thaler R. (1985). Does the Stock Market Overreact? The Journal of Finance, 40, 793-805.

De Fusco, R., Ivanov, S., and Karels, G. (2011). The exchange traded funds' pricing deviation: Analysis and forecasts. Journal of Economics and Finance, 35, 181-197.

Delmas, M. A., Etzion, D., and Nairn-Birch, N. (2013). Triangulating Environmental Performance: What Do Corporate Social Responsibility Ratings Really Capture? Academy of Management Perspectives, 27(3), 255–267.

DeMiguel, V., Garlappi, L. and Uppal, R. (2009). Optimal versus Naive Diversification: How inefficient Is the 1/n Portfolio Strategy? Review of Financial Studies, 22, 1915-1953.

Demsetz, H. and Lehn, K. (1985). The Structure of Corporate Ownership: Causes and Consquences. Journal of Political Economy, 93, 1155-1177.

Derwall, J. Koedijk, K. and Ter H. J. (2011). A tale of values-driven and profit-seeking social investors. Journal of Banking & Finance, Elsevier, 35(8), 2137-2147.

Derwall, J., Guenster, N., Bauer, R. and Koedijk, K. (2005). The Eco-Efficiency Premium Puzzle. Financial Analysts Journal, 61(2), 51–63.

Devore, J. L. (2011). Probability and Statistics for Engineering and the Sciences (8th ed.). Boston, MA: Cengage Learning.

Dimson, E., Marsh, P. and Staunton, M. (2020). Credit Suisse Global Investment Returns Yearbook 2020.

Drenovak, M., Uroševiæ, B., and Jelic, R. (2014). European Bond ETFs: Tracking Errors and the Sovereign Debt Crisis. European Financial Management, 20(5), 958-994.

Dumitrescu, A., Järvinen, J., and Zakriya, M. 2023. Hidden Gem or Fool's Gold: Can Passive ESG ETFs Outperform the Benchmarks? International Review of Financial Analysis. 86 p. 102540.

Eakins, S. G. and Mishkin, F. S. (2012). Financial markets and institutions: global edition 7th ed. (7th). London: Pearson Education Limited.

Eccles, R., Kastrapeli, M. and Potter, S. (2017). How to Integrate ESG into Investment Decision-Making: Results of a Global Survey of Institutional Investors. Journal of Applied Corporate Finance, 29, 125-133.

Economist Intelligence Unit (2015). The cost of inaction: Recognising the value at risk from climate change. Available at: https://eiuperspectives.economist.com/sustainability/cost-inaction.

Edmans, A. (2010). Does the Stock Market Fully Value Intangibles? Employee Satisfaction and Equity Prices. Journal of Financial Economics, 101, 621-640.

Edmans, A. (2012). The Link Between Job Satisfaction and Firm Value, with Implications for Corporate Social Responsibility. Academy of Management Perspectives, 26(4), 1-19.

Elia, M. (2012). Tracking Error of Traditional and Synthetic European Exchange-Traded Funds. SSRN Electronic Journal. Available at: 10.2139/ssrn.2003976.

Ellen MacArthur Foundation (2019). What is a circular economy? Available at: https://ellenmacarthurfoundation.org/topics/circular-economy-introduction/overview.

Ellis, Ch. (1985). Investment Policy: How to Win the Loser’s Game. Homewood, IL: Dow Jones-Irwin.

Ellsberg, D. (1961). Risk, Ambiguity, and the Savage Axioms. The Quarterly Journal of Economics, Oxford University Press, 75(4), 643–669.

Elton, E. J. and Gruber, M. J. (1977). Risk Reduction and Portfolio Size: An Analytic Solution. Journal of Business, 50, 415–437.

Elton, E. J., and Gruber, M. J. (2013). Mutual Funds. In Constantinides, G. M. Harris, M. and Stultz, R. M. (Ed.) Financial Markets and Asset Pricing: Handbook of Economics and Finance, 2 (Part B), Elsevier.

Elton, E. J., Gruber, M. J., Brown, S. J., and Goetzmann, W. N. (2014). Modern portfolio theory and investment analysis. New York: John Wiley & Sons.

Elton, E. J., Gruber, M. J., Comer, G., and Li, K. (2002). Spiders: Where Are the Bugs? The Journal of Business, 75(3), 453–472.

Elton, E. J., Gruber, M. J., Das, S. and Hlavka, M. (1993). Efficiency with Costly Information: A Reinterpretation of Evidence form Managed Portfolios. Review of Financial Studies, 6, 1-22.

Engle, R. F., Sarkar, D. (2006). Premiums-Discounts and Exchange Traded Funds. The Journal of Derivatives, 13(4), 27- 45.

European Securities and Markets Authority. (2024). *Final Report: Guidelines on funds’ names using ESG or sustainability-related terms* (ESMA34-472-440). Retrieved from https://www.esma.europa.eu/sites/default/files/2024-05/ESMA34-472-440_Final_Report_Guidelines_on_funds_names.pdf

ESMA (2022). Performance and Costs of EU Retail Investment Products. ESMA Annual Statistical Report.

ESMA (2023). ESMA Market Report. Costs and Performance of EU Retail Investment Products 2023.
Estrada, J. (2008). Fundamental indexation and international diversification. Journal of Portfolio Management, 34(3), 93-109.
European Commission. (2023). Revision of EU rules on sustainable finance disclosure (SFDR). Retrieved from: https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/14666-Revision-of-EU-rules-on-sustainable-finance-disclosure_en.
European Fund and Asset Management Association (2021). Market Insights. ESG investing in the UCITS market, a powerful and inexorable trend, 4.
Fama, E. F. (1965). Random walks in stock market prices. Financial Analysts Journal 21(5), 55–59.
Fama, E. F. (1965). The behavior of stock-market prices. Journal of Business, 38(1), 34–105.
Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. The Journal of Finance, 25(2), 383–417.
Fama, E. F. and Blume, M. (1966). Filter Rule and Stock Market Trading Profits. Journal of Business, 39, 226-241.
Fama, E. F. and French, K. R. (1993). Common Risk Factors in the Returns on Stocks and Bonds. Journal of Financial Economics, 33, 3-56.
Fama, E. F. and French, K. R. (2015). A Five-Factor Asset Pricing Model. Journal of Financial Economics, 116, 1-22.
Fama, E. F. and French, K. R. (2018). Choosing factors. Journal of Financial Economics, 128, 234–252.
Fama, E. F. and Jensen, M. C. (1983). Separation of Ownership and Control. Journal of Law and Economics, 26, 301- 325.
Fama, E. F., Fisher, L., Jensen, M. C. and Roll, R. (1969). The adjustment of stock prices to new information. International Economic Review, 10(1), 1–21.
Farinha, J. (2003). Corporate governance: a survey of the literature. Review of Financial Markets, 6 (1 & 2), 5-52.
Ferri, R. (2006). All About Index Funds. McGraw-Hill.
Ferri, R. (2009). The ETF Book: All You Need to Know About Exchange-Traded Funds, Updated Edition. John Wiley & Sons.
Fichtner, J., Heemskerk, E., and Garcia-Bernardo, J. (2017). Hidden power of the Big Three? Passive index funds, re-concentration of corporate ownership, and new financial risk. Business and Politics, 19(2), 298-326.
Filippou, I., Gozluklu, A. E. and Rozental, H. (2022). ETF Arbitrage and International Diversification. WBS Finance Group Research Paper, 260, AFA 2022 Annual Meeting, Available at SSRN: https://ssrn.com/abstract=3287417.
Forbes, K. J. and Rigobon, R. (2002). No Contagion, Only Interdependence: Measuring Stock Market Comovements. The Journal of Finance, 57(5), 2223–2261.
Fortin, R. and Michelson, S. (2002). Indexing Versus Active Mutual Fund Management. Journal of Financial Planning, 15, 82–94.
French, K. R. (2008). Presidential Address: The Cost of Active Investing. Journal of Finance, American Finance Association, 63(4), 1537-1573.
Frino, A. and Gallagher, D. (2001) Tracking S&P 500 index funds. Journal of Portfolio Management, 28, 44–55.
Frino, A. and Gallagher, D. (2002) Is index performance achievable? An analysis of Australian equity index funds. Abacus, 38, 200–214.
FTSE Russel (2018). FTSE ESG Index Series. Integrate ESG considerations into a broad market benchmark.
FTSE Russell (2017). Indexes and benchmarks made clear.

FTSE Russell (2017). Insights. The anatomy of smart beta.
Gaba, A. and Kumar, R. (2021). Tracking Error and Pricing Efficiency of Exchange Traded Funds: A Systematic Literature Review. Orissa Journal of Commerce, 42 (3), 42-58.
Gallagher, D. R. and Segara, R. (2005). The performance and trading characteristics of exchange-traded funds. Journal of Investment Strategy, 1 (1), 47-58.
Gastineau, G. L. (2010). The Exchange-Traded Funds Manual. John Wiley & Sons.
Gibson, G. (1889). The Stock Markets of London, Paris and New York, New York: G.P. Putnam's Sons.
Global Asset Management Inc. (2017). Understanding ETF bid-ask spreads.
Global Impact Investing Network (2019). Annual Impact Investor Survey.
Global Reporting Initiative and Sustainability Accounting Standards Board (2021). A Practical Guide to Sustainability Reporting Using GRI and SASB Standards.
Glode, V. (2011). Why mutual funds "underperform". Journal of Financial Economics, 99(3), 546-559.
Golub, B., Novick, B., Madhavan, A., Shapiro, I., Walters, K. and Ferconi, M. (2013). Exchange Traded Products: Overview, Benefits and Myths. BlackRock Investment Institute.
Gompers, P. A., Ishii, J. L. and Metrick, A. (2003). Corporate Governance and Equity Prices. Quarterly Journal of Economics, 118 (1), 107-155.
Grim, D. and Berkowitz, D. (2020). ESG, SRI, and Impact Investing: A Primer for Decision-Making. Vanguard Research.
Grubel, H. G. (1968). Internationally Diversified Portfolios: Welfare Gains and Capital Flows. The American Economic Review, 58, 1299-1314.
Halbritter, G. and Dorfleitner, G. (2015). The Wages of Social Responsibility –Where Are They? A Critical Review of ESG Investing. Review of Financial Economics, 26, 25-35.
Hallez, E. (2022). Passive ESG: More active than the average index fund? Available at: https://www.investmentnews.com/passive-esg-more-active-than-the-average-index-fund-221210.
Harper, J. T., Madura, J. and Schnusenberg, O. (2006). Performance comparison between exchange-traded funds and closed-end country funds. Journal of International Financial Markets, Institutions and Money, 16, (2), 104-122.
Henisz, W., Koller, T., and Nuttall, R. (2019). Five ways that ESG creates value. McKinsey Quarterly, November.
Henriksson, R. D. (1984). Market Timing and Mutual Fund Performance: An Empirical Investigation. The Journal of Business, University of Chicago Press, 57(1), 73-96.
Hill, J. (2020). Environmental, Social, and Governance (ESG) Investing. Cambridge, MA: Academic Press.
Hoepner, A. G. and Zeume, S. Fiduciary Duty and Sin Stocks: Is Vice Really Nice? (2013). Handbook of Institutional Investment and Fiduciary Duty. Available at SSRN: https://ssrn.com/abstract=1485846.
Hong, H. and Kacperczyk, M. (2009). The price of sin: The effects of social norms on markets. Journal of Financial Economics, Elsevier, 93(1), 15-36.
Hsu, J. and Campollo, C. (2006). An Examination of Fundamental Indexation. New Frontiers in Index Investing: Journal of Indexes, 58, 32-37.
Hsu, J. Liu, X., Shen, K., Viswanathan, V. and Zhao, Y. (2018). Outperformance through Investing in ESG in Need. The Journal of Index Investing, 9, 18-26.
https://etfgi.com/news/press-releases/2023/01/etfgi-reports-actively-managed-etfs-listed-globally-gathered-net
https://wildershares.com/about.php

https://www.indexindustry.org/sixth-annual-index-industry-association-benchmark-survey-reveals-continuing-record-breaking-esg-growth-multi-asset-expansion-by-index-providers-globally%EF%BF%BC/
https://www.ssga.com/us/en/intermediary/etfs/resources/education/how-etfs-are-created-and-redeemed.
Hvidkjær, S. (2017). ESG Investing: A Literature Review. Dansif (September).
Ilhan-Nas, T., Okan, T., Tatoglu, E., Demirbag, M., Wood, G. and Glaister, K. W. (2018). Board composition, family ownership, institutional distance and the foreign equity ownership strategies of Turkish MNEs. Journal of World Business, 53 (6), 862-879.
Index Industry Association (2022). 2022 ESG Survey Report.
International Monetary Fund. (2004). Compilation guide on financial soundness indicators (Appendix VII: Glossary). Washington, DC: International Monetary Fund. Retrieved from https://www.imf.org/external/np/sta/fsi/eng/2004/guide/index.htm
Intergovernmental Panel on Climate Change (2018). Annex IV: Expert Reviewers of the IPCC Special Report on Global Warming of 1.5 °C.
Investment Leaders Group (2014). The Value of Responsible Investment. University of Cambridge Institute for Sustainability Leadership. Available at: ILG The Value of Responsible Investment | Cambridge Institute for Sustainability Leadership (CISL).
Ippolito, R. A. (1989). Efficiency with Costly Information: A Study of Mutual Fund Performance: 1965-1984. Quarterly Journal of Economics, 104, 1-23.
Jaffe, J. F. (1974). Special Information and Insider Trading. The Journal of Business, 47, 410-428.
Jane Street (2018). Market Structure and ETF Strategy Important Considerations for NAV-Based ETF Trading.
Jares, T. and Lavin, A. (2004). Japan and Hong Kong Exchange-Traded Funds (ETFs): Discounts, Returns, and Trading Strategies. Journal of Financial Services Research, 25(1), 57-69.
Jensen, M. C. (1968). The Performance of Mutual Funds in the Period 1945-1964. The Journal of Finance, 23, 389-416.
Johnson Jr., Stout, J. and Walter, A. (2020). Profound Change: The Evolution of ESG. Business Lawyer, 75, 2567-2608.
Johnson, B., Bioy, H., Kellett, A., and Davidson, L. (2013). On the right track: Measuring tracking efficiency in ETFs. Journal of Index Investing, 4(3), 35–41.
Johnson, W. F. (2009). Tracking errors of exchange traded funds. Journal of Asset Management, 10(4), 253-262.
Jorion, P. (2004). Portfolio Optimization with Tracking-Error Constraints. Financial Analysts Journal, 59 (5).
Kahneman, D. and Tversky, A. (1979). Prospect theory: an analysis of decision under risk, Econometrica, 47, 263–291.
Kamalnath, A. (2018). Gender Diversity as the Antidote to 'Groupthink' on Corporate Boards. Deakin Law Review 22, 85-106.
Kaserer, C. and Pfau, S. (1993). Perfomance deutscher Aktienfonds, Die Bank, 10, 596-600.
Keeley, B. (2015). Income Inequality: The Gap between Rich and Poor, OECD Insights, OECD Publishing, Paris.
Kempf, A. and Osthoff, P. (2007). The Effect of Socially Responsible Investing on Portfolio Performance. European Financial Management, 13(5), 908-922.
Kendall, M. G. (1953). The analysis of economic time-series—Part I: Prices. Journal of the Royal Statistical Society, series A (General), 116(1), 11–25.
Keynes, J. M. (1923). Some aspects of commodity markets, Manchester Guardian Commercial: European Reconstruction Series, 13, 784–786.

Knight, F. (1921). Risk, uncertainty and profit. Boston: Houghton.
Kon, S. J. and Jen, F. C. (1978). Estimation of Time-Varying Systematic Risk and Performance for Mutual Fund Portfolios: An Application of Switching Regression. The Journal of Finance, 33(2), 457–475.
Kosowski, R. (2011). Do Mutual Funds Perform When it Matters Most to Investors? US Mutual Fund Performance and Risk in Recessions and Expansions. Quarterly Journal of Finance, 1(03), 607-664.
Kotsantonis, S. and Serafeim, G. (2020). Human capital and the future of work: implications for investors and ESG integration. Journal of Financial Transformation, Capco Institute, 51, 115-130.
Kumar, R., Dayaramani, N. and Rocha, J. (2019). Understanding & Comparing ESG Terminology. A Practical Framework for Identifying the ESG Strategy That Is Right for You. State Street Global Advisors.
Lansing, T. (2013). Regulation of exchange traded funds as part of ESMA guidelines implementation: Operational costs and requirements for exchanges. Journal of Securities Operations & Custody, 6(1), 19–24.
La Porta, R., Lopez-de-Silanes, F. and Shleifer, A. (2002). Government Ownership of Banks. Journal of Finance, 57, 265-302.
Larson, A. B. (1960). Measurement of a random process in futures prices. Food Research Institute Studies 1(3), 313– 324.
Lehmann, B. N. and Modest, D. M. (1987). Mutual Fund Performance Evaluation: A Comparison of Benchmark Comparisons. Journal of Finance, 42, 233-265.
Lev, B., Sarath, B. and Sougiannis, T. (2005). R&D Reporting Biases and Their Consequences. Contemporary Accounting Research, John Wiley & Sons, 22 (4), 977-1026.
Lewellyn, D. (2016). Simple Product, Complex Process in: ETF Report. The Magazine for ETF Investors. ETF.com, May 22-23.
Lintner, J. (1965). The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. The Review of Economics and Statistics, 47(1), 13–37.
Lo, A. (2016). What is an index? The Journal of Portfolio Management, 42(2), 21-36.
Lopes, L. (1987). Between hope and fear: The psychology of risk. Advances in Experimental Social Psychology, 20, 255–295.
Lorie, J. H. and Niederhoffer, V. (1968). Predictive and Statistical Properties of Insider Trading. The Journal of Law and Economics, 11 (1), 35-53.
Malkiel, B. (2003). The Efficient Market Hypothesis and Its Critics. The Journal of Economic Perspectives, 17, 59-82.
Madhavan, A. (2016). Exchange-Traded Funds and the New Dynamics of Investing. Oxford University Press.
Malkiel, B. (2011). A Random Walk down Wall Street. New York: W. W. Norton & Company.
Malkiel, B. G. (1973). A random walk down Wall Street (12th ed.). New York: W.W. Norton.
Malkiel, B. G. (2003). A random walk down Wall Street: the time-tested strategy for successful investing. New York: W.W. Norton.
Malkiel, B., Mullainathan S. and Stangle, B. (2005). Market Efficiency versus Behavioral Finance. Journal of Applied Corporate Finance, Morgan Stanley, 17(3), 124-136.
Malladi, R. and Fabozzi, F. (2017). Equal-weighted strategy: Why it outperforms value-weighted strategies? Theory and evidence. Journal of Asset Management, 18(3), 188-208..
Markowitz, H. (1952). Portfolio Selection. The Journal of Finance, 7, 77–91.
Markowitz, H. (1959). Portfolio selection; efficient diversification of investments. New York: Wiley.

Margolis, J. D., Elfenbein, H. A., & Walsh, J. P. (2009). Does it pay to be good... and does it matter? A meta-analysis of the relationship between corporate social and financial performance (Working Paper). Harvard Business School.
Merton, R.C. (1987). A Simple Model of Capital Market Equilibrium with Incomplete Information. Journal of Finance, 42, 483-510.
Meulbroek, L. K. (1992). An Empirical Analysis of Illegal Insider Trading. Journal of Finance, 47(5), 1661-1699.
Meze-Hausken, E. (2000). Migration caused by climate change: how vulnerable are people inn dryland areas? Mitigation and Adaptation Strategies for Global Change, Springer, 5(4), 379-406.
Milonas, N. T. and Rompotis, G. G. (2006). Investigating European ETFs: The Case of the Swiss Exchange Traded Funds. The Annual Conference of HFAA, Thessaloniki, Greece.
Miziołek, T. and Feder-Sempach (2019). Tracking ability of exchange-traded funds. Evidence from Emerging Markets Equity ETFs. Bank i Kredyt, 50 (3): 221-248.
Miziołek, T., Feder-Sempach, E. and Zaremba, A. (2020). International Equity Exchange-Traded Funds. Navigating Global ETF Market Opportunities and Risks. Palgrave Macmillan.
Morningstar (2012). Securities Lending in Physical Replication ETFs: A Review of Providers' Practices.
Morningstar (2021). 2020 U.S. Fund Fee Study.
Morningstar (2021). Spotlight on Synthetic ETFs in Europe.
Moskowitz, T. J. (2000). Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses: Discussion. Journal of Finance, 55: 1695–1703.
Mossin, J. (1966). Equilibrium in a Capital Asset Market. Econometrica, 34(4), 768–783.
MSCI (2018). Introducing ESG Investing. MSCI ESG Research.
MSCI (2020). MSCI ESG Leaders Indexes Methodology.
MSCI. (2022). Foundations of ESG investing: How ESG affects equity valuation, risk and performance.
MSCI (2023). MSCI KLD 400 Social Index (USD) Factsheet.
MSCI (2023). Enhancements to MSCI's Fund ESG Ratings Methodology
Munich Re (2020). Record hurricane season and major wildfires – The natural disaster figures for 2020. Available at: https://www.munichre.com/en/company/media-relations/media-information-and-corporate-news/media-information/2021/2020-natural-disasters-balance.html.
Nagy, Z., Kassam, A. and Lee, L. (2016). Can ESG Add Alpha? An Analysis of ESG Tilt and Momentum Strategies. The Journal of Investing, 25. 113-124.
OECD (2023). OECD Guidelines for Multinational Enterprises on Responsible Business Conduct, OECD Publishing, Paris.
Oman, C. (2001). Corporate Governance and National Development. OECD Development Centre Working Papers, 180, OECD Publishing.
Ormiston, J., Charlton, K., Donald, M. S., and Seymour, R. G. (2015). Overcoming the Challenges of Impact Investing: Insights from Leading Investors. Journal of Social Entrepreneurship, 6 (3), 352-378.
Osterhoff, F. and Kaserer, C. (2016). Determinants of tracking error in German ETFs – the role of market liquidity. Managerial Finance, 42, 417-437.
O'Sullivan, A. and Sheffrin, S. M. (2003). Economics: Principles in Action. Upper Saddle River, New Jersey: Pearson Prentice Hall.
Otten R. and Schweitzer M. (2002). A Comparison between the European and the US Mutual Fund Industry. Managerial Finance, 28, 14-35.

Parks, R. W., Zivot, E. (2006). Financial market efficiency and its implications. University of Washington. Investment, Capital and Finance. Available at: http://faculty.washington.edu/ezivot/econ422/Market%20Efficiency%20EZ.pdf [Accessed 20 May 2022].
Pástor, L., Stambaugh, R. and Taylor, L. (2021). Sustainable investing in equilibrium. Journal of Financial Economics, 142(2), 550-571.
Pastor, L., Stambaugh, R. and Taylor, L. A. (2021). Sustainable investing in equilibrium. Journal of Financial Economics, 42 (2), 550-571.
Pedersen, L. H., Fitzgibbons, S., & Pomorski, L. (2021). Responsible Investing: The ESG-efficient Frontier. Journal of Financial Economics, 142(2), 572–597.
Pettit, R. R. and Venkatesh, P.C. (1995). Insider Trading and Long-Run Return Performance. Financial Management, 24(2), 88-103.
Phengpis, Ch. and Swanson, P. E. (2009). iShares and the US Market Risk Exposure. Journal of Business Finance & Accounting, 36(7-8), 972-986.
Plagge, J. and Grim, D. (2020). Have Investors Paid a Performance Price? Examining the Behavior of ESG Equity Funds. Journal of Portfolio Management, 46 (3), 123–140.
Plyakha, Y., Uppal, R. and Vilkov, G. (2012). Why Does an Equal-Weighted Portfolio Outperform Value- and Price-Weighted Portfolios? SSRN Electronic Journal. Available at: https://ssrn.com/abstract=2724535.
Pope, P. F. and Yadav, P. K. (1994). Discovering Errors in Tracking Error. The Journal of Portfolio Management, 20 (2), 27-32.
Qadan, M. and Yagil, V. (2012). On the dynamics of tracking indices by exchange traded funds in the presence of high volatility. Managerial Finance, 38 (9), 804-832.
Ratanapakorn, O. and Sharma, S. C. (2002). Interrelationships among regional stock indices. Review of Financial Economics, 11(2), 91-108.
Reilly, F. K. and Brown, K. C. (2012). Investment analysis & portfolio management. Mason, OH: South-Western Cengage Learning.
Rendleman, R., Jones, C., Latané, H. (1982). Empirical Anomalies Based on Unexpected Earnings and the Importance of Risk Adjustments. Journal of Financial Economics, 10, 269–87.
Richey, G. (2017). Fewer reasons to sin: a five-factor investigation of vice stock returns. Managerial Finance, 43 (9), 1016-1033.
Robbins, L. (1932). An Essay on the Nature and Significance of Economic Science. Macmillan, London.
Roberts, H. (1967). Statistical versus clinical prediction of the stock market. Unpublished manuscript.
Roll, R. (1977). A Critique of the Asset Pricing Theory's Tests' Part I: On Past and Potential Testability of the Theory. Journal of Financial Economics, 4, 129-176.
Roll, R. (1992). A Mean/Variance Analysis of Tracking Error. The Journal of Portfolio Management, 18, 13-22.
Roll, R. and Ross, S. A. (1984). The Arbitrage Pricing Theory Approach to Strategic Portfolio Planning. Financial Analysts Journal, 40(3), 14–26.
Rompotis, G. (2023). The Compliance of ESG equity funds in the U.S. International Journal of Business and Economic Studies, 5(1), 31-48.
Rompotis, G. G. (2005). An Empirical Comparing Investigation on Exchange Traded Funds and Index Funds Performance. SSRN Electronic Journal. Available at: https://ssrn.com/abstract=903110.
Rompotis, G. G. (2009). A Comprehensive Study on the Seasonality of Greek Equity Funds Performance. South-Eastern Europe Journal of Economics, 7(2), 229-255.

Rompotis, G. G. (2011). Predictable Patterns in ETF's Return and Tracking Error. (2011). Studies in Economics and Finance, 28 (1), 14-35.
Rosenberg, B., Reid, K. and Lanstein, R. (1985) Persuasive Evidence of Market Inefficiency. Journal of Portfolio Management, 11, 9-17.
Ross, S. A. (1976). The Arbitrage Theory of Capital Asset Pricing. Journal of Economic Theory, 13, 341-360.
Rouwenhorst, K., G. (2004). The origins of mutual funds. Yale International Center for Finance Working Paper Series.
Rowley, J., Hirt, J. and Wang, H. (2018). Setting the record straight: Truths about indexing. Vanguard Investments, Canada.
Roy, A. D. (1952). Safety first and the holding of assets. Econometrica, 20, 431–49.
S&P Dow Jones Indices (2022). SPIVA U.S. Scorecard Year-End 2021.
S&P Dow Jones Indices (2023). Dow Jones Sustainability World Index Factsheet.
S&P Dow Jones Indices (2023). S&P ESG Index Series Methodology.
S&P Global (2020). What is the difference between ESG investing and socially responsible investing? Available at: https://www.spglobal.com/en/research-insights/articles/what-is-the-difference-between-esg-investing-and-socially-responsible-investing.
Samuelson, P. A. (1965). Proof that properly anticipated prices fluctuate randomly. Industrial Management Review, 6(2), 41–49.
Scherer, B. (1994). Timing deutscher Investmentfonds. Jahrbuch für Nationalökonomie und Statistik, 213, 187–208.
Schoenfeld, S. (2004). Active Index Investing. Hooboken, New Jersey: Wiley.
Schyra, A. (2013). Indices as Benchmarks in the Portfolio Management. Springer Gabler.
SEC (2012). Investor Bulletin: Exchange-Traded Funds (ETFs).
Securities and Exchange Commission (2014). Investor Bulletin. How Fees and Expenses Affect Your Investment Portfolio.
Sharpe, W. F. (1963). A Simplified Model for Portfolio Analysis. Management Science, 9, 277-293.
Sharpe, W. F. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. Journal of Finance, 19, 425–442.
Sharpe, W. F. (1966) Mutual fund performance. The Journal of Business, 39, 119-138.
Sharpe, W. F. (1991). The Arithmetic of Active Management. Financial Analysts Journal, 47, 7-9.
Sharpe, W. F. (1992). Asset Allocation: Management Style and Performance Measurement. Journal of Portfolio Management, 18, 7-19.
Shefrin, H. and Statman, M. (2000) Behavioral Portfolio Theory. Journal of Financial and Quantitative Analysis, 35(2), 127-151.
Shiller, R. (2013). Sharing Nobel honors, and agreeing to disagree. The New York Times, 26 October. Available at: http://www.nytimes.com/2013/10/27/business/sharing-nobel-honors-and-agreeingto-disagree.html?hp [Accessed 10 May 2022].
Shiller, R. J. (1980). Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends? The American Economic Review, 71(3), 421-436.
Shin, S. and Soydemir, G., (2010). Exchange-Traded Funds, Persistence in Tracking Errors and Information Dissemination. Journal of Multinational Financial Management, 20, 4-5.
Shukla, R. (2004). The value of active portfolio management. Journal of Economics and Business, 56(4), 331-346.
Siegel, L. (2003). Benchmarks and Investment Management. The Research Foundation of AIMR.
Smith, A. (1776). The Wealth of Nations: An Inquiry into the Nature and Causes of the Wealth of Nations.

Staal, A. (2006). Essays in empirical finance. Doctoral dissertation. Northwestern University.
State Street (2023). 2023 Outlook for ETFs. Innovation and Growth Continue.
State Street Global Advisors. (2025). How SPY reinvented investing: The story of the first U.S. ETF. Retrieved from https://www.ssga.com/us/en/intermediary/insights/how-spy-reinvented-investing-story-of-first-us-etf
Statman, M. and Glushkov, D. (2009). The Wages of Social Responsibility. Financial Analysts Journal, 65(4), 33–46.
Stockholm Resilience Centre (2017). Annual Report. Stockholm University, Sweden.
Sushko, V. and Turner, G. (2018). The implications of passive investing for securities markets. BIS Quarterly Review, 113-129.
Svetina, M. and Wahal, S. (2008). Exchange Traded Funds: Performance and Competition. SSRN Electronic Journal. Available at: https://ssrn.com/abstract=1303643.
Swalm, R. D. (1966). Utility theory – insights into risk taking. Harvard Business Review, 47, 123–136.
Swedroe, L. and Adams, S. (2022). Your Essential Guide to Sustainable Investing. Harriman House.
Tobin, J. (1958). Liquidity Preference as Behavior Towards Risk. Review of Economic Studies, 25 (2), 65-86.
Toronto Stock Exchange. (2025, March 6). Toronto Stock Exchange celebrates 35 years of ETF innovation. TMX Group. Retrieved from https://investors.tmx.com/English/news-and-events/news-details/2025/Toronto-Stock-Exchange-Celebrates-35-Years-of-ETF-Innovation/default.aspx
Treynor, J. L. (1965). How to Rate Management of Investment Funds. Harvard Business Review, 43, 63-75.
Tse, Y. and Martinez, V. (2007). Price discovery and informational efficiency of international iShares funds. Global Finance Journal, 18(1), 1-15.
United Nations General Assembly (2015). Transforming our world: the 2030 Agenda for Sustainable Development, A/RES/70/1.
United Nations Environmental Program Finance Initiative (2004). The Materiality of Social, Environmental, and Corporate Governance Issues to Equity Pricing.
United Nations Environment Programme (2015). Paris Agreement. Available at: https://wedocs.unep.org/20.500.11822/20830.
United Nations General Assembly (1948). The Universal Declaration of Human Rights (UDHR). New York: United Nations General Assembly.
United Nations General Assembly (2015). Transforming our world: the 2030 Agenda for Sustainable Development. A/RES/70/1.
United Nations Secretary-General (2016). Report of the Inter-Agency and Expert Group on Sustainable Development Goal Indicators. E/CN.3/2016/2/Rev.1.
US SIF Foundation (2020). Report on US Sustainable and Impact Investing Trends.
Von Neumann, J., and Morgenstern, O. (2007). Theory of games and economic behavior. Princeton University Press.
Von Wyss, R. (2004). Measuring and predicting liquidity in the stock market. PhD dissertation, University of St. Gallen.
Weiner, A., Sun, X. and Banerjee, S. (2020). ESG Index Construction A Tool to Build Sustainable Capital Markets. State Street Global Advisors.
Whelan, T., Atz, U., Van Holt, T. and Clark, C. (2021). ESG and financial performance. Uncovering the Relationship by Aggregating Evidence from 1,000 Plus Studies Published between 2015 – 2020, Centre for Sustainable Business, NYU-Stern, 1, 2015-2020.
White, H. (1980). A Heteroskedasticity-consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. Econometrica. 48(4), p. 817.

Winegarden W., (2019). ENVIRONMENTAL, SOCIAL, AND GOVERNANCE (ESG) INVESTING: An Evaluation of the Evidence. Pacific Research Institute. Available at: https://www.pacificresearch.org/wpcontent/uploads/2019/05/ESG_Funds_F_web.pdf.

WisdomTree (2021). What is an ETF and How Does it Differ from a Mutual Fund?

Womack, K. L. (1996). Do Brokerage Analysts' Recommendations Have Investment Value? Journal of Finance, 51(1), 137-167.

Wong, K. H. Y. and Shum, W. C. (2010). Exchange-traded funds in bullish and bearish markets. Applied Economics Letters, 17:16, 1615-1624.

Wooldridge, J. M. (2010). Econometric Analysis of Cross Section and Panel Data (2nd ed.). Cambridge, MA: MIT Press.

Working, H. (1934). A random-difference series for use in the analysis of time series. Journal of the American Statistical Association, 29(185), 11–24.

World Economic Forum (2022). Defining the 'G' in ESG Governance Factors at the Heart of Sustainable Business.

Yan, T. and Qian, W. Y. (2004). Environmental migration and sustainable development in the upper reaches of the Yangtze River. Population and Environment, 25(6), 613–636.

Yu, E., Luu, B. and Chen, C. (2020). Greenwashing in environmental, social and governance disclosures. Research in International Business and Finance, Elsevier, 52(C).

Zerbib, O. (2022). A sustainable capital asset pricing model (S-CAPM): Evidence from environmental integration and sin stock exclusion. Review of Finance, 26 (6), 1345-1388

LIST OF TABLES

Table 1. Percentage of U.S. equity funds underperforming benchmarks 36
Table 2. Classification of security weighting methods.. 47
Table 3. Division of financial market indices by type.. 49
Table 4. Classification of security selection approaches in index construction 50
Table 5. Sustainable Development Goals with selected indicators 71
Table 6. Examples of ESG equity indices by rating providers.. 84
Table 7. Key differences between ETFs and mutual funds 90
Table 8. Average annual ongoing costs of EU UCITS products (in %)................................ 104
Table 9. ETF Investment Universe 108
Table 10. Direct causes of the ETF tracking difference 114
Table 11. Summary of inclusion and exclusion criteria in the research sample 121
Table 12. Explanation of variables used in the study 124
Table 13. Summary of research steps and applied methods... 126
Table 14. Regional exposure of ETFs in the research sample.. 130
Table 15. Descriptive statistics of ETF assets under management.. 131
Table 16. Descriptive statistics of ETF age 131
Table 17. Descriptive statistics of ETF total expense ratios... 132
Table 18. Distribution of ETF replication methods.. 133
Table 19. Descriptive statistics of benchmark risk 133
Table 20. Results of normality and homogeneity tests for independent variables 133
Table 21. Significance of differences in independent variables.. 134
Table 22. Descriptive statistics of tracking errors (%) for ESG and non-ESG ETFs............ 137
Table 23. Normality and homogeneity assumptions for dependent variables....................... 138
Table 24. Significance of differences between tracking errors of ESG and Non-ESG ETFs 139
Table 25. Descriptive statistics of variables used in the model (full ETF sample) 141
Table 26. Correlation matrix for the full sample of ETFs 141
Table 27. Results of the Levin–Lin–Chu unit root test (full sample).................................... 143
Table 28. Results of the ADF test (full sample) 143
Table 29. Comparison of Pooled OLS, RE, and FE models (full sample)............................ 143
Table 30. Diagnostic tests for model selection... 144
Table 31. Diagnostic tests for the Fixed Effects model (full ETF sample) 145

Table 32. Tests for endogeneity and multicollinearity in the Fixed Effects model (full sample) 146
Table 33. Specification of the linear dynamic panel model (full ETF sample)..................... 147
Table 34. Evaluation of lagged TE_2 as an instrumental variable in the GMM model (full sample) 147
Table 35. Estimation results of the GMM model (full ETF sample).................................... 148
Table 36. Diagnostic tests of the GMM model (full ETF sample)....................................... 149
Table 37. Robustness checks for the GMM model (full ETF sample)................................. 150
Table 38. Descriptive statistics of continuous variables for ESG and non-ESG ETFs 150
Table 39. ADF test results for ESG and non-ESG ETFs.. 151
Table 40. Diagnostic tests for model selection in ESG and non-ESG ETF samples 151
Table 41. Diagnostic tests for Random Effects models in ESG and non-ESG ETFs............ 152
Table 42. Tests for endogeneity and multicollinearity in the Random Effects models (ESG and non-ESG ETFs) 153
Table 43. Evaluation of lagged TE_2 as an instrumental variable in the GMM model (ESG ETFs) 154
Table 44. Evaluation of lagged TE_2 as an instrumental variable in the GMM model (non-ESG ETFs) 154
Table 45. Estimation results of the GMM model (ESG ETFs) 155
Table 46. Estimation results of the GMM model (non-ESG ETFs)..................................... 155
Table 47. Diagnostic tests of the GMM model (ESG and non-ESG ETFs).......................... 156
Table 48. Chow test results for GMM models (ESG vs. non-ESG ETFs) 157
Table 49. GMM model estimation with ESG interaction terms (full ETF sample) 157
Table 50. Diagnostic tests of the model with ESG interaction terms (full ETF sample) 158

LIST OF FIGURES

Figure 1. Relationship between risk and return 14
Figure 2. The efficient frontier for alternative portfolios 17
Figure 3. Characteristic security line 19
Figure 4. Capital market line 22
Figure 5. Security market line 23
Figure 6. Portfolio value from investing USD 100,000 over 20 years 37
Figure 7. Average Expense Ratios of Actively Managed and Index Mutual Funds in the U.S. Equity and Bond Markets, 1996–2021 38
Figure 8. Total Assets Under Management in U.S. Passive and Active Mutual Funds and ETFs, and Passive Share of the Market (1995–2019) 40
Figure 9. Sustainable Investing Spectrum 55
Figure 10. Scope of governmental factors 67
Figure 11. Operation of Exchange-Traded Funds 93
Figure 12. Primary information disclosed by exchange-traded funds 105
Figure 13. Sample Selection Process 122
Figure 14. Growth in ESG ETF Assets and Number of Funds (2006-2024) 123
Figure 15. Distribution of total expense ratio (TER) for ESG and Non-ESG ETFs 134
Figure 16. Distribution of assets under management (AUM) for ESG and Non-ESG ETFs 135
Figure 17. Distribution of age for ESG and Non-ESG ETFs 135
Figure 18. Distribution of benchmark risk (RISK) for ESG and Non-ESG ETFs 136
Figure 19. Correlation heatmap of variables under study 136
Figure 20. Correlation heatmap of ETFs' tracking errors 138
Figure 21. Distribution of tracking error (TE_1) for ESG and Non-ESG ETFs 139
Figure 22. Distribution of tracking error (TE_2) for ESG and Non-ESG ETFs 139
Figure 23. Distribution of tracking error (TE_3) for ESG and Non-ESG ETFs 140
Figure 24. Dependencies between variables for the full sample of ETFs 142
Figure 25. Autocorrelation function (ACF) for TE_2 for the full sample of ETFs 145
Figure 26. Autocorrelation Function (ACF) for TE_2 of ESG and Non-ESG ETFs 152