ESG-Integrated machine learning portfolio opimization strategies: performance analysis and applications

Auteur : Dalne, Grégoire
Promoteur(s) : Ittoo, Ashwin
Faculté : HEC-Ecole de gestion de l'Université de Liège
Diplôme : Master en ingénieur de gestion, à finalité spécialisée en Financial Engineering
Année académique : 2022-2023
URI/URL : http://hdl.handle.net/2268.2/17534

ESG-INTEGRATED MACHINE LEARNING PORTFOLIO OPTIMIZATION STRATEGIES: PERFORMANCE ANALYSIS AND APPLICATIONS

Jury:
Supervisor:
Ashwin ITOO
Reader:
Julien HAMBUCKERS

Master thesis presented by
Grégoire DALNE
For a Master's degree in Business Engineering, specialization in Financial Engineering
Academic year 2022/2023

Acknowledgments

First and foremost, I would like to express my sincere gratitude for the guidance provided by my supervisor, Professor Ashwin Ittoo, during the preparation of this thesis. His flexibility and precious advice made him a great mentor to me and have been critical to my learning process and the successful completion of this thesis.

Furthermore, I extend my gratitude to Professor Julien Hambuckers, who served as the primary reader for this master thesis. His spontaneous and valuable suggestions have significantly improved the quality of this work.

Sincere thanks are extended to the administration, faculty, and staff of HEC Liège. Their dedicated work and efforts to foster a virtuous learning environment equipped me with the academic rigor and professional values necessary for a successful career in finance.

I also want to express my appreciation to the people who gave me the unique chance to turn my internships at Deloitte and Degroof Petercam into successful experiences. These experiences have offered me a real-world perspective that has enriched my understanding of the financial industry and provided me with skills that turned out invaluable to write this master thesis.

Lastly, I want to thank my family and friends. Their unconditional support throughout my years at HEC have been vital and helped me make the best out of my academic journey.

Table of content

Table of figures

1. Introduction

1.1. Context and main objective of the thesis

In recent years, portfolio managers and finance practitioners have experienced an increasing need to adopt quantitative methods, surpassing any previous levels, as part of their goal to achieve superior returns. The range of these methods spans over a wide spectrum, encompassing basic linear models that lay the foundation for quantitative analysis, all the way to complex deep-learning networks that hold immense potential to revolutionize financial decision-making. This increasing dependence on quantitative approaches can be attributed to a selection of factors that have profoundly influenced and shaped the modern financial landscape.

Recent advancements in technology and computational power have revolutionized portfolio management by enabling efficient processing and analysis of vast amounts of data. This has created opportunities to develop sophisticated models that capture complex market patterns. Moreover, the availability of historical financial data has expanded, allowing practitioners to comprehensively back-test and validate investment strategies, refining their portfolio management approaches based on empirical evidence. Furthermore, the acknowledgment of behavioral biases in investment decision-making has forced practitioners to adopt objective and data-driven methods. Quantitative models offer a systematic and disciplined approach to mitigating these biases, promoting more consistent and rational investment decisions.

D'Hondt et al. (2020) have shown in their paper that AI-managed portfolios can provide significant gains for low-education/income investors among a sample of Belgian investors. However, these results are mixed if we consider all investor profiles, and the realized gain compared to a passive strategy does not appear significant on average. These conclusions raise questions about the effectiveness of machine learning approaches in portfolio management as compared to more simple methodologies and highlights the need for further investigation and refinement of these models and their implementation.

The research conducted in this master thesis will go beyond those conclusions by enlarging the investment universe, which was restricted to each investor's securities for which he/she made a transaction in the past two years before the study occurred. By considering a broader definition of the range of accessible assets and a more comparable investment universe from year to year, our study aims to provide a more comprehensive understanding of the performance of quantitative models in different market conditions and across various investment options.

In addition to constraints related to risk aversion, the matter of preferences in terms of the sustainability of the investment will be explored. With the increasing focus on environmental, social, and governance (ESG) factors in investment decision-making, integrating sustainability considerations in the portfolio optimization process has become a crucial aspect of modern portfolio management. This master thesis aims to explore how asset allocation suggested by quantitative models can incorporate sustainability metrics and preferences to construct portfolios that align with investors' ESG goals.

In this master thesis, we aim to explore various regression models with consideration of the recommendations and findings from the scientific literature. To determine the most effective approach, we will describe and investigate different machine learning models and their possible implementation for portfolio optimization, including advanced regression techniques and deep learning networks. These models have shown promising results in the context of portfolio optimization, as highlighted in the scientific literature. By leveraging these techniques, we aim to access to improved predictive power and ultimately improve the performance of portfolio management strategies.

Our primary objective is to compare diverse range of regression models, by examining the performance of their respective investment strategies within a defined investment universe. Through a comprehensive out-of-sample comparison, we aim to evaluate the returns generated by each strategy, ultimately identifying the ones that exhibit the potential to significantly enhance investment performance. Also, in the final stages of the portfolio construction process we will suggest, we will actively investigate ESG consideration. By integrating these crucial factors into our decision-making process, we aim to create portfolios that not only yield attractive returns but also promote positive societal and environmental impact.

1.2. Research motivations

The BCG Global Asset Management Market Sizing 2022 study (McIntyre et al, 2022) shows that the global assets under management have grown from USD 31 trillion in 2003 to USD 112 trillion in 2021, at an accelerating growth rate. These figures include all assets professionally managed in exchange for management fees. The asset management industry plays a crucial role in the global financial ecosystem, providing solutions to the investment needs of individuals, institutions, and corporations.

According to the same study, in 2020, two-thirds of these assets were actively managed, whereas one-third were passively managed (with methods such as index tracking). While passive strategies offer simplicity and low costs, active strategies aim to generate excess returns by leveraging market insights, research capabilities, and quantitative models. Given the substantial portion of assets actively managed, it is evident that portfolio managers face immense pressure to deliver superior performance and justify their fees.

Thousands of investment firms are sharing this market, and it is thus crucial for each of them to exhibit better performance in comparison to competitors to stand out and attract investors. In an industry driven by competition, the need to generate alpha has never been so tense. Therefore, implementing quantitative methods that could help portfolio managers to reach superior performance is of crucial importance and the stakes are high.

These professionals are also very attentive to academic breakthroughs in the finance and statistical fields, as an investing strategy can lose in performance if it is widely adopted. It has been shown that the public knowledge of the factors identified in scientific literature impacts their ex-post predictive quality (McLean & Pontiff, 2016). Consequently, identifying and implementing new methods for portfolio optimization is key in the asset management industry.

In parallel, ESG considerations have become increasingly important in asset management, as investors and regulators recognize the significance of environmental, social, and governance factors in investment decision-making. Asset managers are now integrating ESG criteria into their investment processes to assess the long-term sustainability and risk profile of their investments. In this context, regulatory frameworks like the Markets in Financial Instruments Directive II (European Union, 2014) play a vital role. Implemented in 2018, MiFID II requires asset managers to disclose detailed information about their investment products, including the potential impact of sustainability factors like ESG considerations on investment decisions.

1.3. Contributions

The primary goal of this master thesis is to comprehensively investigate various quantitative methods, with a particular focus on deep learning models, used by practitioners or proposed by academic research. The aim is to measure the extent to which these methods can be useful when combined to achieve superior performance compared to passive strategies such as index trackers. To achieve this objective, the thesis will address the following research questions:

1. Among a selection of quantitative methods, determine the one which promotes the best asset allocation in a pre-defined investment universe, that yields the best out-of-sample portfolio performance based on the mean-variance optimization model.

2. For one of the discussed regression models, determine what happens if we introduce ESG considerations. We want to understand the return cost or benefit of ESG preferences at asset allocation stage and compare it with the realized cost or benefit of ESG preferences and understand the predictability of this measure.

By addressing these research questions, this master thesis aims to provide valuable insights and empirical evidence to guide portfolio managers and investment professionals in their pursuit of superior performance and effective portfolio allocation strategies by describing a comprehensive approach that could be used as a reference for sustainable portfolio management in the future.

1.4. Structure of the thesis and general methodology

The literature review will provide an overview of research on sustainable portfolio management and machine learning methods for portfolio optimization. The literature review discusses different portfolio management styles and emphasizes the importance of selecting appropriate optimization techniques, in particular the Black-Litterman or Markovitz models, to maximize wealth and address criticisms against Socially Responsible Investing (SRI) strategies. It also explores the application of machine learning models in financial markets to enhance return predictions.

In parallel, the literature review will explore strategies for integrating ESG considerations into portfolio construction. Firstly, it will describe the research conducted by Pedersen et al. (2021), which introduces the notion of an ESG-efficient frontier. This study provides valuable insights into incorporating environmental, social, and governance factors into portfolio management. Furthermore, the review will analyze a more recent model proposed by Steuer and Utz (2023). Their model introduces the concepts of non-contour efficient fronts and an efficient surface, which seek to incorporate ESG preferences into the process of constructing investment portfolios.

The methodology section begins by stating the research objectives of the study. It then describes the data sources and programming languages used in the research, including the selection of source databases and the programming language utilized. The methodology also explains the selection of macroeconomic variables and investment universe for the study, as well as the considerations regarding the chosen time window.

The next part of the methodology provides a detailed overview of the linear methods used for expected return estimation, including Ridge Regression, Lasso Regression, and Elastic Net Regression. A comparison and practical considerations of these methods are also discussed. The methodology further delves into non-linear methods such as Multilayer Perceptrons (MLP), Recurrent Neural Networks (RNNs), and Long Short-Term Memory (LSTM) networks, explaining their working principles and their implementation in investment strategies.

The concluding segment of the methodology focuses on the integration of ESG considerations during the portfolio construction phase. This section heavily draws upon the insights and principles elucidated by Steuer and Utz (2023) and Pedersen et al. (2021). Its primary objective is to measure the impact, in terms of return, associated with varying degrees of ESG integration. By quantifying the costs related to different levels of ESG integration, this analysis seeks to provide valuable insights into the trade-offs involved in incorporating ESG factors into portfolio decision-making.

2. Literature review

In this section, we will mainly investigate the current state of research on sustainable portfolio management, but also on machine learning methods that can be implemented to achieve portfolio optimization. We will thus consider these two aspects in two separate sub-sections that will be summarized in a short concluding section describing how we can combine the findings from the two main sections of this literature review.

2.1. Current state of research in sustainable portfolio management

This area of study has gained significant momentum in recent years due to the increasing awareness of the social, environmental, and governance issues affecting financial markets. We will thus go through the main concepts and empirical findings on sustainable investing for the sake of giving context to our study. Then, we will cover two promising models that may have a considerable impact on investment decisions in the future, which will be implemented further in this thesis with real market data.

2.1.1 Sustainable investing: definition and empirical performance

In their book, Marszk and Lechman (2023, p58) define the main feature of sustainable investing (investment) as the fact of "*taking into consideration ESG factors in the investment decisions including but not limited to the selection of securities (or other assets), portfolio management and any associated research*".

In terms of performance, we may wonder whether investing only in socially responsible securities is a constraint that represents a burden in our portfolio optimization problem, or if it represents a winning factor. Numerous studies have tried to compare market performance and sustainable investment funds' performance across various sectors or geographical areas, yielding mixed results (Pedersen et al., 2021). Consequently, we will reference one of the most recent and comprehensive studies available about this topic at the time of writing, as it effectively builds upon the findings of preceding research.

In their study, Dumitrescu et al. (2023) have examined the performance of 121 passive U.S. equity SRI ETFs between January 2010 and December 2020, comparing it to a benchmark portfolio comprising passive S&P 500 ETFs. The findings drawn for ETFs, a type of investment representing a basket of selected securities such as stocks or bonds, can easily be extended to portfolio management and as such, we can confidently draw some interesting conclusions from this study.

The research reveals that SRI ETFs do not consistently underperform their passive non-SRI benchmark ETFs. In the second half of the sample period (2015-2020), which corresponds to the industry's development and growth, there was no statistically significant difference between the performance of SRI ETFs and their benchmarks. Interestingly, investors could have potentially gained significant value from SRI ETFs in the last two years of the sample (2019-2020). These results remain unaffected even when accounting for the effect of fees on SRI ETFs' returns.

These findings have important implications for portfolio managers, as incorporating ESG preferences in investment decisions could help investors outperform both markets and passive benchmark ETFs. The study also provides insights on the strategies (inclusion or exclusion) that drive the under- or over-performance of SRI ETFs compared with their benchmarks. It is found that only the SRI ETFs incorporating inclusion (positive screening) and, specifically, those employing the environmental inclusion strategy are drivers of abnormal returns.

Dumitrescu et al. (2023) have thus shown that good financial performance and sustainable attributes of investments are not always mutually exclusive, and they conclude that SRI ETFs can offer transparent and cost-efficient investment opportunities with significant potential to contribute to sustainability and sustainable development. They also conclude that for investors looking to integrate social responsibility and sustainability into their investment choices, financial underperformance may not always be a fatality, especially when selecting investments that incorporate an environmental inclusion strategy which has shown to perform best in the last two years of the study's sample.

2.1.2. Portfolio management styles regarding sustainable investments

We have previously discussed that investors have an increasing interest in investing their money into greener financial assets. Now, we will try to understand how asset managers have been effectively meeting this demand, and why there is room for improvement. In this literature review, we will make with a comprehensive description of the approach suggested by Steuer and Utz (2023) for sustainable portfolio management. We will also review the model described by Pedersen et al. (2021), which uses a different though interesting approach to tackle the same problem.

In their paper, Steuer and Utz (2023) start by providing a description of the current methodology used by asset managers for portfolio construction both if they consider ESG preferences of their clients or not. A distinction is also made between portfolio construction approaches that make use of a simple ESG screening from those that go further by also implementing a tri-criterion optimization model, which we will comment on in-depth as it will be the base of our study. Steuer and Utz (2023) distinguish three categories of mutual funds type, that can be regarded as asset allocation styles in the context of this master thesis.

2.1.2.1 Conventional mutual funds.

This type of mutual funds refers to traditional funds, for whom ESG considerations are not considered at any step in the portfolio construction process. They basically start by defining the investment universe. This process is defined as *Screening*. Stocks, bonds, funds, or any other security can be selected based on their historical returns track record, as well as their expected performance and future risk, among other criteria such as availability, liquidity, or tax considerations.

Then, the mutual fund must find a tradeoff between risk and return. The method used at this stage is to pick a portfolio that lies on the efficient frontier, initially described by Markowitz (1952), as part of his Modern Portfolio Theory. Along this frontier, every point shows an optimal portfolio that offers the highest expected return for a given level of risk, or the lowest risk for a given expected return. The goal is then to pick the portfolio along the efficient frontier that maximizes the ratio between return and risk, that is often measured with help of the Sharpe ratio. The selection of the optimal portfolio also depends on the investor's risk tolerance and investment objectives.

In summary, conventional mutual funds focus solely on risk and return criteria when constructing portfolios, without considering ESG factors. This approach, based on the efficient frontier model, has been the standard for portfolio management for decades. However, as investors become more aware of the environmental, social, and governance impacts of their investments, there is a growing demand for more sustainable investment strategies, which has led to the development of ESG-screened and tri-criterion optimization approaches.

2.1.2.2 Non-integrated ESG mutual funds.

Non-integrated ESG mutual funds, as discussed by Steuer and Utz (2023), differ from conventional mutual funds in that they incorporate an additional ESG screening step in the portfolio construction process. This screening is typically applied during the first stage, where the investment universe is defined. Securities that do not meet a minimum ESG score or do not comply with specific ESG criteria are excluded from the investment universe. However, as demonstrated by Utz et al. (2014), after the ESG screening, the portfolio construction process follows the same methodology as conventional mutual funds, using the efficient frontier model and focusing only on finding the optimal risk-return trade-off.

This approach is referred to as "negative screening" or "exclusionary screening" by the authors, as it removes companies or assets that do not meet the ESG criteria, rather than actively selecting those with strong ESG performance. As already discussed, it has been shown empirically that exclusion strategies perform poorer as compared to inclusion strategies (Dumitrescu et al., 2023).

While non-integrated ESG mutual funds can provide investors with a more sustainable investment option compared to conventional mutual funds, they do not optimize the risk-return-ESG tradeoff. As a result, serious ESG investors might find this approach insufficient for their sustainability objectives.

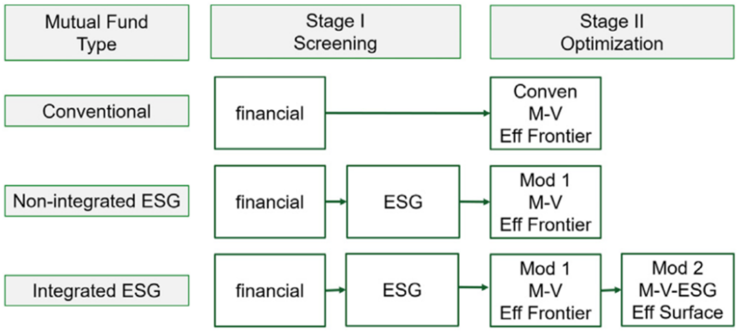
2.1.2.3 ESG-integrated mutual funds.

Integrated ESG mutual funds, on the other hand, go beyond simple ESG screening by incorporating ESG factors into the portfolio optimization process. Steuer and Utz (2023) propose a tri-criterion optimization model that accounts for risk, return, and ESG performance for portfolio construction. This approach results in an "M-V-ESG efficient surface" which represents the set of optimal portfolios that offer the best possible trade-offs between risk, return, and ESG performance, depending on the investors preferences.

Unlike non-integrated ESG mutual funds, the integrated ESG approach actively selects securities based on their ESG performance, in addition to risk and return considerations. This allows investors to better align their investments with their sustainability objectives and potentially achieve superior long-term risk-ESG-adjusted returns. In addition, as discussed above, implementing an inclusion strategy has shown to be driving better financial performance as compared to exclusion strategies (Dumitrescu et al., 2023).

According to Steuer and Utz (2023), ESG-integrated mutual funds represent a more advanced and comprehensive approach to sustainable portfolio management compared to non-integrated ESG mutual funds. By incorporating ESG factors into the portfolio optimization process and utilizing a tri-criterion optimization model, these funds can better meet the needs of serious ESG investors who seek to optimize the risk-return-ESG tradeoff (Figure 1).

Figure 1.
Mutual Funds Types with regard to ESG Criterions Integration.



Note. Source: Steuer, R.E. and Utz, S. (2023). Non-contour efficient fronts for identifying most preferred portfolios in sustainability investing. European Journal of Operational Research, 306 (2), 742-753.

2.1.3. What optimization model should we consider for sustainable investments?

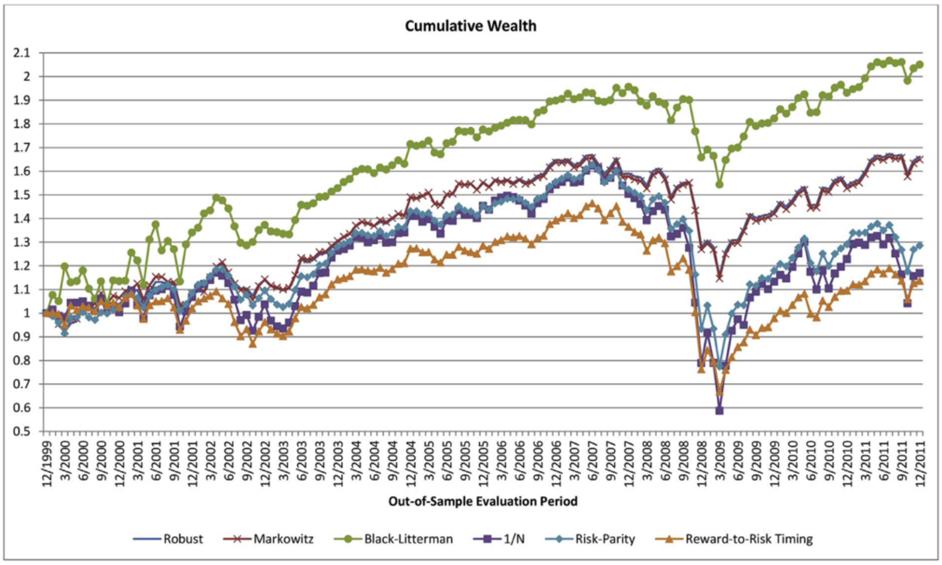
Oikonomou et al. (2017) highlighted the significance of the optimization process within the framework of Socially Responsible Investing. Their research, based on a significant sample of US stocks, showed that the choice of optimization technique has a profound impact on SRI fund management, with sophisticated quantitative optimization methods found to lead to maximized wealth on a risk-adjusted basis, even after accounting for transaction costs. They further discovered that such methods offer more significant advantages for SRI portfolios compared to those portfolios constructed from an unscreened sample of stocks.

The research underlined that the optimization process matters, indicating the importance of not just the type of "responsible" screening criteria used, but also the optimization technique employed in determining optimal investment weights. The findings advocate for the application of more sophisticated optimization techniques, like the Markovitz (1952) and Black and Litterman (1992) models, which contradict much of the mainstream finance literature that supports naïve or simplistic techniques.

According to Oikonomou et al. (2017), the Black-Litterman model (1992) usually emerged as the best technique, while naïve diversification was usually the worst, especially in terms of risk, risk-adjusted returns, diversification, and intertemporal stability of the SRI portfolios. The Mean-variance optimization model suggested by Markovitz (1952) appeared to be the second-best choice, with a clear margin as compared to more naïve methods. The authors also highlight that, in addition to robust performance as compared to more naïve methods, the application of these quantitative optimization techniques can indeed answer some of the criticisms against SRI strategies, describing them as too qualitative and subjective.

Considering this, the research by Oikonomou et al. (2017) reinforces the importance of examining the Markovitz and Black-Litterman models within the context of this master's thesis on sustainable investments. This thesis will however primarily focus on the mean-variance optimization model of Markovitz (1952) and related models as the core methodology. Indeed, two recent ground-breaking models in terms of ESG investing were recently presented by Steuer et al. (2023) and Pedersen et al. (2021), relying mostly on the initial concept of the efficient frontier, and it has been chosen to investigate these two models extensively. While these two sophisticated models only will be relied upon for ESG-integrated portfolio construction, further exploration of the Black-Litterman model and the application of machine learning techniques to it for portfolio construction may provide fertile ground for additional research, based on the results presented by Oikonomou et al. (2017) (Figure 2). In the two next sub-sections, we will describe more in-depth the findings of the two optimization models specific to sustainable investing that we just mentioned.

Figure 2.
Cumulative Wealth for various Optimization Methods in the context of SRI Investment.



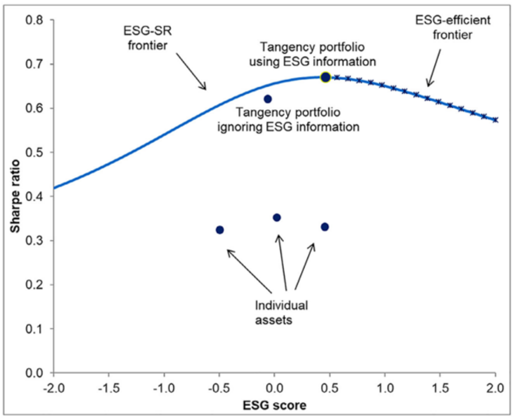
Note. Source: Oikonomou, I., Platanakis, E., & Sutcliffe, C. (2017). Socially responsible investment portfolios: Does the optimization process matter? British Accounting Review, 50(4), 379–401.

2.1.4. Efficient frontier between ESG score and Sharpe ratio.

In their groundbreaking paper, Pedersen et al. (2021) introduced the concept of the ESG-efficient frontier, which illustrates the highest achievable Sharpe ratio for each normalized ESG level (Figure 3).

The primary objective of this concept is to understand the final costs or benefits associated with ESG investing. By utilizing this model, investors can select a portfolio with the highest attainable Sharpe ratio based on their ESG preferences. A subsequent comparison can then be made between the characteristics of this portfolio and those suggested by the traditional mean-variance optimization model, providing insights into the costs or benefits of incorporating ESG considerations into investment decisions.

Figure 3.
ESG-Efficient Frontier (Pedersen et al., 2021).



Note. Source: Pedersen, L.H., Fitzgibbons, S. & Pomorski L. (2021). Responsible investing: The ESG-efficient frontier. Journal of Financial Economics, 142(2), 572-597

While this concept has paved the way for further research, it has also attracted criticism from other scholars. Since its publication, new research has emerged, proposing an alternative model that offers a more comprehensive assessment of the implications of ESG investing.

2.1.5. Efficient surface and non-contour efficient fronts

Steuer and Utz (2023) have investigated a novel approach to sustainable investing in their study. They suggest that modern investors no longer consider finding their most preferred risk/return tradeoff on the mean-variance efficient frontier introduced by Markovitz (1952) and are increasingly interested in finding their preferred risk/return/ESG tradeoff, indirectly on the mean-variance-ESG efficient surface.

In their paper, the authors define a serious ESG investor as "*an investor whose strength of interest in ESG makes that consideration a criterion competitive on the playing field of portfolio selection with risk and return, thus causing the investor's efficient frontier to become an efficient surface*" (Steuer & Utz, 2023, p.742).

The authors argue that mutual funds don't achieve the maximum that could be done to propose an asset allocation that would be in line with the ESG preferences of these serious ESG investors as defined above, and that at the time of writing the mutual fund industry does not have the necessary knowledge to fully satisfy this type of investors. Indeed, solving this novel tri-fold optimization problem can be a real challenge for asset managers. In that regard, the authors propose the novel concept of non-contour efficient curves that are stretched over the whole M-V-ESG efficient surface to help for better understanding and portfolio selection based on a given investor's risk aversion and ESG preferences characteristics.

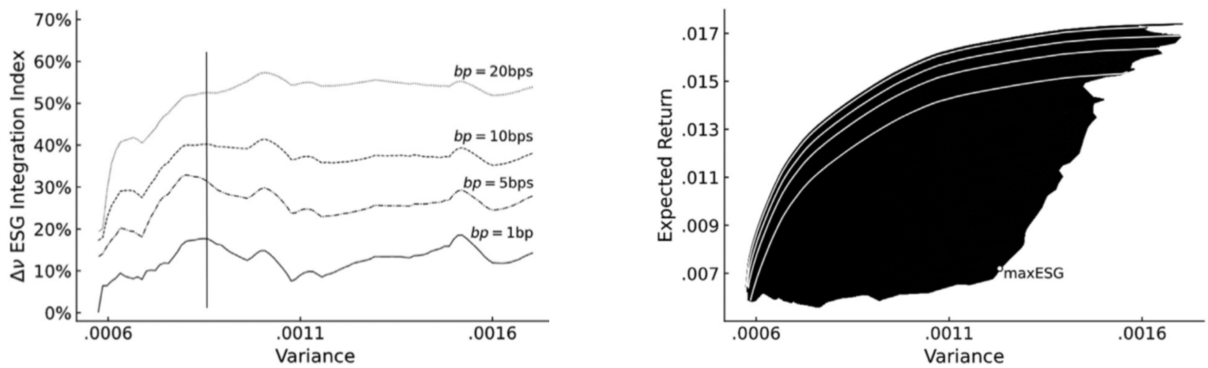
The bi-criterion solution presented by Pedersen et al. (2021) as an approach to address the same problem has been subject to criticism in the research conducted by Steuer and Utz (2023). According to Steuer and Utz (2023). While the conversion of the tri-criterion portfolio optimization problem faced by serious ESG investors into a bi-criterion problem enhances comprehension and implementation, it does possess a significant drawback.

Specifically, the method described by Pedersen et al. (2021) fails to account for the risk aversion exhibited by investors, despite attempting to strike an optimal balance between portfolio Sharpe ratios and ESG scores. They state that the pursuit of attaining the highest Sharpe ratio achievable based on an investor's ESG preferences does not guarantee the selection of low-variance portfolios for risk-averse individuals. Likewise, it does not ensure the allocation of investments towards more aggressive strategies for those who are willing to assume greater risks to achieve higher returns.

The implementation of the mean-variance-ESG efficient surface model suggested by the authors allows to handle these shortcomings. The concept of efficient surface, however, might be difficult to grasp. In their paper, Steuer and Utz (2023) suggest a precise methodology to construct and interpret the efficient surface of a particular investment universe.

To do so, they introduce the concept of non-contour efficient fronts (Figure 4). Each non-contour efficient front is characterized by an amount of basis point in terms of return that a serious ESG investor is willing to give up to achieve better ESG portfolio performance. The more a serious ESG investor is willing to give up on return, the greener his/her investment. Each NC-efficient front represents portfolios that provide the best trade-off between return and risk for his specific level of allowed basis point relaxation. By displaying all the possible non-contour efficient fronts among the investment universe side by side, we can thus construct the mean-variance-ESG efficient surface (Figure 4).

Figure 4.
Non-Contour Efficient Fronts (left) and M-V-ESG Efficient Surface (right) – (Steuer & Utz, 2023).



Note. Source: Steuer, R.E. and Utz, S. (2023). Non-contour efficient fronts for identifying most preferred portfolios in sustainability investing. European Journal of Operational Research, 306 (2), 742-753.

In order to measure the increased ESG performance by making a return tradeoff, the authors have developed an interesting approach instead of considering ESG scores only: the ESG integration index.

$$\Delta \upsilon = \frac{\upsilon^T x^{bp,i} - \upsilon^T x^i}{\nu_{max} - \upsilon^T x^i}$$

In this definition, υ^T is the ESG Coefficient Vector, in which each element is an ESG score that corresponds to a security in the investment universe. x^i is the optimal portfolio as suggested by the classical mean-variance optimization model, without consideration of the ESG preferences of the investor. $x^{bp,i}$ is a portfolio for which a basis point quantity bp has been given up by the investor to achiever higher ESG performance. ν_{max} is the portfolio with the highest attainable ESG score. In most cases, this portfolio is indeed an undiversified investment in the security with the highest ESG score in the considered investment universe.

The ESG integration index $\Delta\upsilon$ described by Steuer and Utz (2023) provides several advantages Firstly, it quantifies the degree of ESG integration in a portfolio, allowing investors to assess the extent to which ESG factors are considered. It provides a scalar value that indicates the magnitude of ESG improvements achieved by relaxing expected returns by a certain number of basis points. Secondly, the index can be used in relative terms, accommodating different ESG metrics used by various financial data providers. This allows for meaningful comparisons across different scales and ranges of ESG scores. Additionally, the index enables visual representation, as demonstrated below, showing the relationship between ESG integration and basis point relaxations in expected returns. It reveals that smaller relaxations yield larger gains in ESG integration, while larger relaxations result in diminishing returns.

Now that we understand better the concept of non-contour efficient fronts described by Steuer and Utz (2023), we can explore their step-by-step approach to construct their efficient surface for a particular investment universe and choose a preferred portfolio on it.

1. <u>Perform a classical mean-variance optimization to build the efficient frontier</u>. It involves the calculation of expected returns, and the variance-covariance matrix among all available securities (x^i). The most preferred portfolio p is obtained by maximizing the Sharpe ratio, which is a key metric to describe risk-return trade-off.

2. <u>Finding the ESG coefficient vector</u>. The ESG coefficient vector contains the ESG scores for each of the securities under consideration (υ^T). These scores can be sourced from ESG rating agencies.

3. <u>Constructing the NC-efficient fronts graph</u>. For various bp (basis point) values chosen by the investor, NC-efficient fronts are constructed. The NC-efficient front represents portfolios that provide the best trade-off between return, risk, and ESG score for a particular level of bp relaxation. Each NC-efficient front can thus be represented as a graph showing the different trade-offs between risk (variance) and return for a particular investor with specific ESG preferences, quantified as the maximum bp relaxation tolerated.

4. <u>Evaluating the NC-efficient fronts</u>. Using $\mathcal{V}_0$, the variance of the most preferred portfolio p on the efficient frontier with no ESG considerations as a reference, the investor evaluates the non-contour efficient fronts to find the most suitable NC-front by making a tradeoff between return and ESG preferences only.

5. <u>Selection of the most preferred point</u>. The investor selects his/her most preferred portfolio, most likely the one that maximizes the Sharpe ratio, among those on the non-contour efficient front selected at the previous step. This selection process allows for personal preference, and to account for individual risk tolerance and ESG preferences.

6. <u>Obtaining the optimal portfolio</u>. After taking the inverse image of the selected point (i.e., determining the portfolio weights that result in the selected point), the optimal portfolio for the decision-maker is obtained.

2.2. Current state of research in quantitative methods for portfolio management

As a complement to the previous discussion about the optimization model we should select (see 2.1.3.), the subsequent sections of this thesis will present a comprehensive overview of the necessary parameters needed for the application of the mean-variance optimization model proposed by Markovitz (1952). Primarily, these will involve the estimations of expected return for each stock in our investment universe and the construction of the expected variance-covariance matrix. Both elements are the center input of our portfolio selection model, influencing the optimization process and consequently, the portfolios' risk and return characteristics. As we proceed, we will discuss the specific methods and models employed in the estimation process of these parameters and describe how they integrate the mean-variance optimization model and contribute to a more robust portfolio selection process.

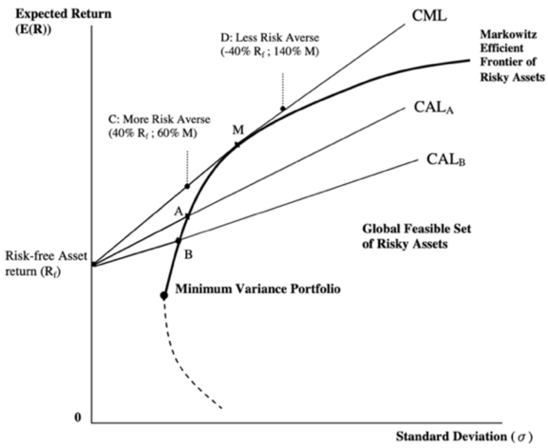
2.2.1. The mean-variance optimization problem

The mean-variance optimization problem, described by Markowitz for the first time in 1952, is described in his Modern Portfolio Theory. It is a quantitative approach that seeks to understand and balance the trade-off between risk and return in an investment portfolio. This approach uses statistical measures, specifically the level of return measured as the expected returns of assets, and the level of risk measured as the portfolio returns expected volatility.

The expected return of a portfolio is calculated by considering all the possible returns that we expect the constituting assets will deliver, weighting them by their respective weights, and then summing them up. The higher the expected return, the better the portfolio is expected to perform in terms of financial returns. The variance, on the other hand, measures the dispersion of the asset's returns and corresponds to the expected standard deviation of returns. It is indeed a measure of portfolio riskiness. The greater the variance, the higher the risk and uncertainty of the portfolio returns in the future.

Markowitz's theory states that by combining different assets with varying expected returns and variances, an investor can construct a portfolio that delivers the maximum possible return for a given level of risk. This concept leads to the creation of what is known as the 'efficient frontier', along which every portfolio has the minimum level of risk for all attainable levels of return, and the maximum return for any possible level of risk. In practice, it can be visualized as a graph that plots the expected returns against the standard deviation, and the resulting curve represents all the optimal portfolios an investor can hold (Figure 5).

Figure 5.
Mean-Variance Efficient Frontier Theoritical Representation (Markovitz, 1952).



Note. Source: Hodnett, K., & Hsieh, H. (2012). Capital Market Theories: Market Efficiency Versus Investor Prospects. International Business & Economics Research Journal, 11(8), 849-862.

In Markovitz's Modern Portfolio Theory, investors then choose portfolios along this efficient frontier based on their specific risk tolerance. Those with high risk tolerance may choose portfolios towards the higher end of the frontier (top-right quadrant), where the expected returns are higher but come with more risk. Conversely, risk-averse investors may choose portfolios towards the lower end of the frontier (bottom-left) quadrant, where the expected returns are more modest, but the associated risk is lower.

This concept revolutionized the field of finance and investment, as it provided a systematic, mathematical way to consider risk and return simultaneously when making investment decisions. It is important to note, however, that the mean-variance optimization approach relies on several assumptions, such as investors being rational and markets being efficient, which do not really hold in the real world.

Having gained a deeper understanding of the Mean-variance optimization problem proposed by Markowitz, we can now address the most demanding aspect of its resolution: determining the most accurate estimations for expected returns and anticipated risk levels to build the efficient frontier. Conventional methodologies depend on the utilization of historical return averages to project expected returns. Similarly, the computation of the standard deviation of these returns is used to estimate the anticipated level of risk.

2.2.2. Future volatility estimation: variance-covariance matrix and more recent developments

The variance-covariance matrix is a pivotal component of the modern portfolio theory, particularly in the context of the mean-variance optimization problem framework. This concept was introduced by Markowitz (1952). Markowitz's framework seeks to optimize a portfolio to maximize expected return for a given level of risk, or equivalently, to minimize risk for a given level of expected return.

The variance-covariance matrix provides a measure of risk associated with a set of assets in a portfolio. Each element in the i-th row and j-th column represents the covariance between the i-th sotck and j-th stock. The elements along the diagonal of the matrix measure the risk associated with each individual asset, namely the variance, while the covariances off the diagonal measure the degree to which two assets move together.

However, the variance-covariance matrix can be highly sensitive to the estimation period and can be unstable over time. To tackle this issue, various methods have been proposed to "shrink" the variance-covariance matrix towards a more stable target. Ledoit and Wolf (2003) introduced a method known as the Ledoit-Wolf shrinkage estimator. This method pulls the estimated covariances towards a structure that is more stable and less prone to estimation error.

Another approach, proposed by Friedman et al. (2008), employs a technique known as the graphical lasso. This method provides a sparse estimate of the inverse covariance matrix, effectively shrinking many off-diagonal elements to zero and therefore leading to a simpler and more interpretable covariance structure. At the time of publishing, the authors estimated that the graphical lasso method was 30 to 4000 times faster than competing methods.

2.2.3. Estimating expected returns using machine learning models.

Several strategies can be employed to forecast future returns. One straightforward approach is using the average of historical returns. However, this seemingly simple method comes with its own set of challenges. While it is intuitive and easy to understand, relying on the average of historical returns as a predictor for future performance can be problematic. This method assumes that past patterns will continue. However, markets are complex and influenced by a multitude of factors such as economic cycles, political events, and technological advancements. Consequently, historical returns may not accurately reflect future possibilities.

Indeed, relying on historical averages often falls short of accurately reflecting the specificities of asset returns, that are observable and have been identified by the finance literature for long. Cont (2001) describes a set of stylized statistical properties of asset returns. He characterizes these properties, also termed 'stylized empirical facts', as empirically observable statistical features of asset returns in the market. In total, he identifies eleven of these stylized statistical properties:

“ - *Absence of autocorrelation (except for small intraday time scales)*
- *Heavy tails*
- *Gain/loss asymmetry*
- *Aggregational Gaussianity*
- *Intermittency*
- *Volatility clustering*
- *Conditional heavy tails*
- *Slow decay of autocorrelation in absolute terms*
- *Leverage effect*
- *Volume/volatility correlation*
- *Asymmetry in time scales” (Cont, 2001, p.224)*

These characteristics highlight that stock returns indeed possess several unique traits that we can capitalize on to enhance the accuracy of predictions, outperforming the method of using past return averages as a metric for forecasting future returns. This is precisely where the recent development of machine learning models proves to be very valuable, because of its unrivaled ability to identify patterns in large datasets to provide with accurate predictions.

Statistical models can enable us to detect diverse patterns in historical data, extending beyond past stock returns to encompass corresponding macroeconomic variables or stock-specific characteristics. This can be done for instance with <u>linear regression</u> models in which we select some variables that we believe have a high explanatory power for predicting stock returns. The most notable studies in that view are most probably:

- Fama and French (1992), in their seminal paper, proposed a three-factor model, extending the simple Capital Asset Pricing Model (CAPM) by including size and book-to-market factors as additional factors to describe past returns and thus predict expected stock returns.

- Carhart (1997) extended the Fama and French three-factor model by adding a fourth factor - momentum - to explain the cross-sectional return patterns in mutual fund performance. This model, often referred to as the Carhart four-factor model, has been widely used to evaluate the performance of mutual funds and other types of portfolios, as the intercept of this model, also referenced as *Jensen's alpha*, can measure how asset managers have been able to generate returns that are not explained by the four factors chosen as independent variables in the linear regression model.

We also have the possibility to construct models that incorporate a multitude of independent variables other than these described by these authors, chosen from a range of macroeconomic or stock-specific factors. To that extent, statistical methods such as Ridge, Lasso, or Elastic Net regressions prove to be extremely beneficial. These techniques are regularization methods designed to enhance the interpretability of the model and prevent overfitting, particularly when dealing with many predictors.

Ridge regression is effective when many predictors contribute to the outcome, as it shrinks the coefficients of less important features without reducing them to zero. On the other hand, Lasso regression can zero out coefficients, thereby performing variable selection when we believe only a subset of predictors influences the output. Elastic Net regression combines these approaches, offering a balance between coefficient shrinkage and variable selection. Thus, these techniques help us focus on the most influential variables, reducing the impact of less important ones, and thereby enhancing the precision and interpretability of our model.

While the current models can perform well and have served as a reference in the financial sector for years, there remains potential for enhancement. It is essential to recognize that the financial markets systematically operate on non-linear dynamics. As evidence to support that claim, we can refer to Barkoulas and Baum (1996), who highlighted the non-linear nature and temporal dependency persistence of stock returns among the contemporary DJIA components.

By leveraging the power of machine learning methods, we can amplify our prediction capabilities, thereby augmenting our predictive accuracy through the ability to identify and exploit non-linear patterns in the input data. A multitude of machine learning models commonly used for times series predictions stands ready to address the inherent non-linear characteristics of financial markets. In the scope of this master thesis, we will focus on three types of machine learning methods to attempt to make better predictions of expected returns among our investment universe and build the efficient frontier more accurately:

- Multilayer Perceptrons (MLP): As an integral type of artificial neural network, MLPs consist of a minimum of three layers of nodes, including input, hidden, and output layers. Their structural design makes them adept for pattern recognition, classification, and even regression tasks, applications of which can be found in various fields including stock return prediction.

- Recurrent Neural Networks (RNN): Specifically designed to recognize patterns in sequential data, RNNs are another category of artificial neural networks. Their unique structure is particularly suited to time series prediction, making them a valuable asset for predicting stock returns, a task that heavily relies on historical and sequential data.

- Long Short-Term Memory (LSTM): The LSTM architecture, a specific variant of recurrent neural network design, has attracted substantial attention within the deep learning community. Its special architecture allows for more robust handling of long-term dependencies in sequence data, which is critical in tasks such as stock return prediction, where recognizing patterns over extended periods is essential.

2.3. Summary

The research areas of sustainable portfolio management and machine learning for portfolio optimization have grown rapidly. This literature review dives into these exciting fields, exploring how they have changed over time, the strategies used, and the hurdles faced when trying to meet both financial and sustainability goals in investment management. This literature review seeks to dissect these complex and evolving fields, focusing on the trends, strategies, and the challenges that investors and fund managers face in achieving financial as well as sustainable objectives.

First, the literature review describes the evolution and performance of sustainable investing. As revealed in a study by Dumitrescu et al. (2023), SRI ETFs do not consistently underperform their non-SRI counterparts. This pattern is especially pronounced when an environmental inclusion strategy is incorporated. This significant finding suggests that investors do not need to sacrifice financial gains for sustainability goals systematically, and that both sustainability and return performance aspects can be beneficially co-existent in investments.

The strategies and approaches in portfolio management are multiple, especially concerning sustainable investments. This review outlines three distinct styles, identified by Steuer and Utz (2023): Conventional mutual funds, which overlook ESG factors; Non-integrated ESG mutual funds, which deliberately exclude poor ESG performers; and ESG-integrated mutual funds, which actively select securities based on their ESG performance. Steuer and Utz (2023) argue in favor of ESG-integrated funds, stating that these provide an optimal risk-return-ESG trade-off, aligning investments more closely with sustainability objectives. Hence, ESG-integrated mutual funds emerge as a promising avenue for investors seeking a balance between financial and sustainability goals.

However, the strategies for sustainable investing can be highly complex. The need for advanced optimization techniques in SRI is paramount, as highlighted by Oikonomou et al. (2017). They recommend sophisticated models like the Markowitz (1952) and Black-Litterman (1992) models over simple methods, asserting their superior capabilities. Pedersen et al. (2021) expand this argument, introducing the ESG-efficient frontier to balance Sharpe ratios and ESG scores, further refining portfolio optimization.

Despite this progress, Steuer and Utz (2023) push for a more nuanced approach with their mean-variance-ESG efficient surface concept. This three-dimensional optimization seeks to address the limitations of the bi-criterion solution proposed by Pedersen et al. (2021). They believe this strategy allows for better accounting of investor risk aversion, proposing non-contour efficient fronts for visualizing and navigating this complex model. This approach could potentially offer a more inclusive evaluation of ESG investments, contributing actively to the ongoing conversation on sustainable investing strategies.

Modern portfolio management extensively utilizes quantitative methods. The mean-variance optimization as proposed by Markowitz (1952), aiming to optimize the trade-offs between risk and return, is considered as one of the most performant and comprehensible model in that regard. This method depends on the accurate estimation of expected return for each stock and the construction of a variance-covariance matrix.

This matrix is at the center of the portfolio selection process, but its accurate estimation poses challenges, particularly in future volatility prediction. Recent developments, such as the Ledoit-Wolf shrinkage estimator (Ledoit & Wolf, 2003) and the graphical lasso (Friedman et al., 2008) have been employed to stabilize the variance-covariance matrix by leveraging named shrinkage methods.

The estimation of expected returns has also seen a shift from reliance on historical return averages to the incorporation of machine learning models. These models, with their ability to identify patterns in large datasets, provide more accurate predictions. Advanced regression models like Ridge, Lasso, or Elastic Net regressions are often used to isolate the most influential variables, thereby reducing the impact of less important ones. It is possible to go further to increase prediction capabilities, by leveraging more sophisticated machine learning methods like Multilayer Perceptrons (MLP), Recurrent Neural Networks (RNN), and Long Short-Term Memory (LSTM) networks.

In summary, the literature presents a vast range of strategies, methods, and techniques for sustainable portfolio management and optimization, providing valuable insights for investors and fund managers. With a clear trend towards incorporating machine learning for enhanced prediction and optimization, and towards the inclusion of ESG considerations, the future of sustainable investing appears promising.

3. Methodology

3.1. Research objectives

This master's thesis is driven by two primary goals. The first one was: "Among a selection of quantitative methods, determine the one which suggests the best asset allocation in a pre-defined investment universe, that yields the best portfolio performance based on the mean-variance optimization problem".

To reach this goal, we will investigate the seven regression models we already discussed in the literature review. We also make the distinction between linear and non-linear regression models. Linear regression models refer to classical regression models, with or without various penalization methods. Non-linear regression models refer to deep learning methods, that are designed to identify non-linear patterns among the provided data.

<u>Linear regression models</u>
- Linear regression
- Ridge regression
- Lasso regression
- Elastic net regression

<u>Non-linear regression models</u>
- Multilayer perceptrons
- Recurrent Neural Networks (RNN)
- Long Short-Term Memory Neural (LSTM) Networks

Each of these models can yield, using a rolling-window approach for training, an expected return for every stock in our pre-defined investment universe at any time step, based on the database constructed as discussed in the next section.

The level of risk, despite the limitations of the classical variance-covariance matrix and the existing solutions to solve these issues already described in the literature review, will be estimated with this measure in our approach. This decision is motivated by two arguments:

- **Computational cost**. Shrinking the variance-covariance matrix, with both linear and non-linear methods, significantly increases the computation time. Also, the fact of considering three distinct variance-covariance matrixes means that, at each time step iteration, we should perform three different Mean-variance optimizations, therefore tripling the overall computation time as compared to a single variance-covariance matrix approach.

- **Clarity.** Our methodology is already quite complex, and it is preferable to base our analysis on a single approach to estimate the level of risk to be able to isolate and thus better understand the impact of each regression approach used to suggest asset allocation.

For the second research objective, which was: "For one of the discussed regression models, determine what happens if we introduce ESG considerations. We want to understand the return cost or benefit of ESG preferences both at asset allocation stage and ex-post stages and compare our results with these of Steuer and Utz (2023)", we will discuss a comprehensive procedure, whose goal is to quantify the return cost to integrate ESG preferences among investment decision-making. We also want to make a link with the first research question and assess to what extent machine learning methods could reduce the cost of ESG preferences integration in portfolio management.

3.2. Data source, programming languages and investment universe selection

In this subsection, we mainly consider practical considerations such as the data providers we will make use of to constitute our working dataset, as well as the choice of the programming languages used for modeling, but also the investment universe that has been selected for our research.

3.2.1. Source databases

For conducting this research, finding reliable and accurate financial and macroeconomic data is vital. In that regard, the choice of Refinitiv Eikon and the Federal Reserve Economic Data (FRED) database as primary data sources appeared as an evidence. These platforms were selected based on their global reputation, extensive coverage, data reliability, and user-friendly nature, all crucial to ensure robust research outcomes.

Refinitiv Eikon was chosen for stock prices data due to its extensive global coverage, the timeliness of its data, and its advanced analytical tools. Known for its accuracy and reliability, Eikon provides real-time and historical data, facilitating a comprehensive understanding of stock price trends. It is indeed one of the most popular financial providers, widely used by finance professionals. Furthermore, its user-friendly interface and its brilliant Excel add-in tool simplifies data collection and initial analyses, proving its value in financial research.

On the other hand, the FRED database was selected for macroeconomic variables due to its comprehensiveness, reliability, and ease of access. This data provider offers a brilliant R API tool to import macroeconomic data easily. Managed by the Federal Reserve Bank of St. Louis, FRED database is considered as a reliable data source and is widely used for economic research around the world.

3.2.2. Programming language

The selection of programming languages and tools in research should align with the specific needs of data analysis and the employed techniques. In this case, the R programming language was predominantly chosen for its robustness, and within the R environment, the TensorFlow and Keras packages were utilized for certain machine learning methods, specifically Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks.

R was primarily selected for its capabilities in statistical analysis and data visualization. It provides a wide range of packages and libraries specifically designed for financial and econometric analyses, making it a suitable choice for handling time series data. Additionally, R's syntax is intuitive, and its community support is substantial, ensuring the availability of solutions to common issues.

However, when it came to implementing RNNs and LSTMs, Python was preferred over R. This choice was driven by Python's superior support for these deep learning techniques, facilitated by the powerful libraries TensorFlow and Keras. These libraries, although primarily associated with Python, can be used within the R environment, providing robust, flexible, and efficient tools for building and training neural networks.

3.2.3. Macroeconomic variables selection

As previously stated, we gathered the macroeconomic variables used as regressors in our prediction models from the Federal Reserve Economic Data (FRED) database website, provided by the Federal Reserve Bank of St. Louis. These variables exclusively pertain to the United States, which is a noteworthy aspect considering that our stock sample consists entirely of US-based companies. The selection of variables has been made in a subjective manner, incorporating a wide range of covariates that were judged to likely correlate with stock returns based on their nature (Appendices, Table 1). This approach aims to maximize the chances for pattern recognition in our models.

3.2.4. Investment universe selection

The investment universe of our study corresponds to the constituting stocks of the Dow Jones Industrial Average (DJIA). As a globally recognized index, the choice of focusing on its constituents confers several advantages for this exploration. Selecting these guarantees that the research is built upon data from firms that have considerable exposure on both the U.S. market and the global economy. These firms are traditionally linked with robust liquidity, reducing the impact of market microstructure noise on our conclusions.

The availability of data stand as another compelling reason for this choice. Premium historical data for these large-cap companies is easily accessible and similar across various data providers, and they are subjects to high standards in terms of financial disclosures. This ensures a high degree of data consistency, vital for comprehensive and accurate statistical analyses.

In addition to the benefits already discussed, the Dow Jones Industrial Average (DJIA) exhibits a strong level of sector diversification (Appendices, Table 2). This enables a more holistic view of the market, as it includes companies from various industries like technology, healthcare, finance, and more. This diversification ensures that the effects of sector-specific shocks are diffused, leading to more reliable analysis results.

Moreover, these 30 companies have significant exposure to the macroeconomic variables contained in our dataset, extracted from the Federal Reserve Economic Data (FRED) database. This relationship enhances the relevance and applicability of our statistical models, as the selected securities are more likely to display meaningful correlations and responses to macroeconomic trends.

However, it is important to recognize the limitations of using the DJIA. One primary concern is the potential bias due to the changing nature of the index. The DJIA is not a static index; its constituents can and do change over time, often reflecting the historical success of the included companies. This means that the stocks under study are typically “winners”, potentially leading to a survivorship bias in our findings.

Moreover, this bias also implies that our research might not be directly extendable to other kinds of markets. Markets with different characteristics or those primarily composed of smaller or less successful companies may not behave in the same way as the DJIA constituents. Therefore, while our findings will provide valuable insights into the behavior of large, successful companies within diverse sectors, caution should be exercised when applying these results to broader or different market contexts.

We should also stress that the 30 stock constituents of the DJIA will indeed not be included in our investment universe sample. Dow, the material science unit of the chemical producer DowDuPont Inc., was separated from its parent company on April 1, 2019 (Reuters Staff, 2023). Consequently, this stock was not quoted before 2019, which is out of our required 20-years time window (from 2001 to 2021) quotation requirement. In addition, Visa Inc. and Salesforce Inc. were not publicly traded before the start of our defined rolling window, therefore it has been decided to not include these stocks in our sample.

3.2.5. ESG rating system

Refinitiv's ESG rating system offers valuable insights to assess the sustainability of financial securities and is among one of the most comprehensive ESG databases in the industry, covering over 85% of the global market cap. Their ESG scores are designed to transparently and objectively measure a company's relative ESG performance, commitment, and effectiveness based on company-reported data. Accessing the ESG scores is made convenient through the Refinitiv Eikon Excel add-in, with a similar extraction procedure as for stock prices. On their website, Refinitv define their scale, which is displayed on Table 3 in the Appendices.

Refinitiv's ESG scores play a crucial role in constructing the ESG Integration Index later in this study. The linear scale from 0 to 100 provided by Refinitiv is well-suited for this purpose, enabling to quantify and compare companies' ESG performance easily. Indeed, by leveraging Refinitiv's data, we can construct an index that captures the integration of ESG factors into investment processes accurately by comparing the aggregate ESG score of a portfolio on a scale ranging from the ESG score of the mean-variance optimal portfolio to the highest achievable ESG score in our investment universe.

3.2.6. Time window considerations

In our analysis, we adopted a time window of 20 years, where we partitioned the data into distinct subsets for training and prediction purposes. Specifically, we utilized a rolling window approach, where 17 years of data were allocated for training our regression models, while the remaining 3 years were reserved for making predictions. By allocating 17 years for training, we aimed to leverage a substantial amount of historical data to capture underlying patterns and relationships among variables and avoid overfitting as we are dealing with over 30 regressors and monthly data. This allowed us to build robust regression models that could effectively learn from past observations and make strong predictions if fed with unseen data.

At each time step, after the training phase, we used the fitted model to predict the next observation of the dataset using a rolling-window approach. This method allowed us to continuously update the model by shifting the time window forward at each step, incorporating the most recent data while excluding the oldest observation. It helped us evaluate the model's performance over multiple time periods and assess its reliability. This approach ensured robust and adaptable forecasts, capturing underlying patterns while considering potential changes in variables' behavior over time.

3.3. Detailed overview of linear methods used for expected return estimation.

Penalization methods have become vital tools for portfolio management, particularly when dealing with high-dimensional data, multicollinearity, or noisy observations that are the standard in finance. We will provide an overview of three popular penalization methods in the context of portfolio management: Ridge, Lasso, and Elastic Net. We discuss the theoretical underpinnings of each technique, their applications in estimating expected returns and shrinking the variance-covariance matrix, and their strengths and weaknesses in the scope of this master thesis.

3.3.1. Ridge Regression

Ridge Regression, proposed by Hoerl and Kennard (1970), is a linear regression method that employs an L2 penalty term. Ridge regression effectively deals with multicollinearity by shrinking the regression coefficients towards zero, thus reducing the variance of the model estimates. In portfolio management, Ridge can be used to stabilize estimates of expected returns in the presence of multicollinearity. However, unlike Lasso, Ridge does not perform variable selection, which may limit its interpretability in high-dimensional settings.

3.3.2. Lasso Regression

Lasso Regression, introduced by Tibshirani (1996), is another linear regression technique that incorporates an L1 penalty term. This penalty term leads to both variable selection and continuous shrinkage, providing sparse solutions that facilitate model interpretability. In the context of portfolio management, Lasso can be applied to estimate expected returns while simultaneously shrinking the variance-covariance matrix, enabling more robust portfolio optimization (Friedman et al., 2008)

3.3.3. Elastic Net Regression

Elastic Net Regression, developed by Zou and Hastie (2005), combines the L1 penalty of Lasso and the L2 penalty of Ridge regression. This hybrid approach addresses the limitations of Lasso and Ridge, particularly in cases with highly correlated features or when the number of predictors exceeds the sample size.

3.3.4. Comparison and Practical Considerations

Each penalization method offers unique advantages and drawbacks for portfolio management. Ridge is more appropriate for addressing multicollinearity without the need for variable selection, while Lasso is well-suited for models requiring sparsity and interpretability, as well as shrinking the variance-covariance matrix. Elastic Net provides a balance between the two, offering both shrinkage and variable selection. Selecting the appropriate method depends on the specific problem, data structure, and the desired trade-off between interpretability and prediction accuracy in portfolio management. In the scope of this master thesis, we will thus examine the performance of each method within one of the two sustainable finance frameworks that we discussed before.

3.4. Detailed overview of non-linear methods used for expected return estimation.

The notations and the ideas presented in the last section are drawn entirely from the authors that are mentioned below. The descriptions given in this section have been included to provide a more comprehensive context for the reader interested in the functionality of artificial neural networks. They are not claimed to be original thoughts, and their formulation should be attributed to the work of the mentioned authors.

3.4.1. Multilayer Perceptrons (MLP)

Multilayer Perceptrons (MLPs) are a specific type of feedforward artificial neural network (ANN). MLPs, as compared to the linear regression models we just discussed, have a highly flexible architecture and can model complex non-linear relationships, making them suitable for time series forecasting, including stock returns prediction. MLPs are trained using a backpropagation algorithm that iteratively adjusts the weights between nodes to minimize the difference between the actual and predicted output. We attempt to provide a more detailed description of the MLPs below by summarizing an extract from Chapter 2 of "Neural Networks and Deep Learning" by Nielsen (2015).

3.4.1.1. How does it work?

Multilayer Perceptrons (MLPs) consist of three layers of interconnected neurons, where each layer processes the output from the previous one, applies a set of weights and biases, and passes the result through an activation function. The output from the final layer forms the network’s overall output. MLPs are said to be *feedforward* networks, meaning that information flows only in one direction, from the input layer to the output layer and passing by the hidden layer(s), without any feedback connections.

MLPs leverage the power of backpropagation, an algorithm introduced by Rumelhart et al. (1986). Backpropagation calculates the gradient of the cost function with respect to the network's weights and biases (Figure 6), thus indicating how changes in these parameters impact the cost. This information guides the adjustment of the weights and biases to minimize the cost function, enhancing the network's learning capability.

3.4.1.2. The cost function

To train a neural network, we need a measure of how well the network is performing. This is traditionally achieved through the usage of a cost function. The cost function quantifies the difference between the predicted output of the network and the true output for a given input. The goal is to minimize this difference, indicating that the network is accurately capturing the underlying patterns in the data.

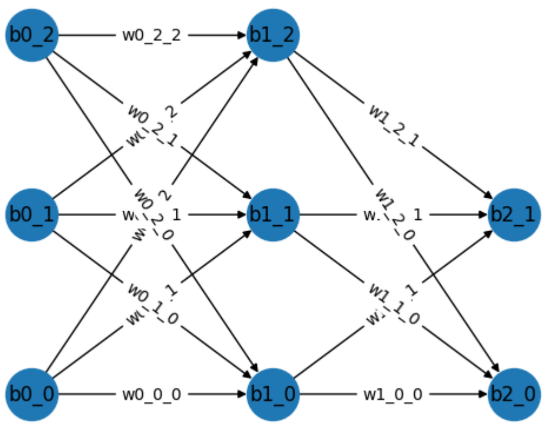
The choice of the cost function depends on the specific task at hand. For example, in a binary classification problem, where the goal is to assign inputs to one of two classes, a common choice is the cross-entropy loss function. In regression problems, where the goal is to predict a continuous value, the mean squared error (MSE) function is often used. For regression, the mean absolute error (MAE) function is similar to the MSE function and may also constitute a good choice, especially for tasks in which there is no necessity to attribute an important weight to large errors.

Given a training example with input x and corresponding true output (label) y, the cost function C is defined as a function of the predicted output $\mathrm{a^L}$, where L represents the index of the output layer. The cost function captures the mismatch between the predicted output and the true output:

$$\mathrm{C = C(a^L, y)}$$

The goal of the training process is to find the set of weights and biases that minimize the overall cost function across all training examples.

Figure 6.
Multilayer Perceptrons Model with 3 input nodes, 2 hidden nodes, and 2 output nodes.



Note. Source: Python networks library

3.4.1.4. Backpropagation algorithm.

Backpropagation, as described by LeCun et al. (2015), is a foundational algorithm used in the training of multilayer perceptrons but also other neural networks. According to the authors, it functions by applying the chain rule for derivatives in a unique way that efficiently calculates the gradient of an objective function in relation to the weights throughout the network. This "backwards propagation" of computation starts at the output layer of the network and progresses towards the input layer, allowing for the determination of gradients at all layers. Once these gradients are established, the corresponding adjustments to the weights of each module can be made, facilitating the learning process within the network. LeCun et al. (2015) state that this transformative approach was first discovered in the 1970s and 1980s and initially met with skepticism. It has however proven to be effective at avoiding issues of local minima in large networks. This groundbreaking algorithm is at the heart of deep learning and has been instrumental for various artificial intelligence applications.

3.4.1.5. Multilayer perceptrons and their implementation in investment strategies

Multilayer Perceptrons, which are indeed a specific implementation of Artificial Neural Networks (ANN), have been established as a powerful tool for predicting financial market trends. This capability was demonstrated by the comparative analysis conducted by Maciel and Ballini (2009) on North American, European, and Brazilian Stock Market Indexes. In their study, the authors stress that the robustness of ANNs is exhibited in their ability to generalize from noisy data, enabling them to predict unseen trends.

This results in a forecasting model that is not only more accurate than the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model, which was used as a benchmark in that study, but is also capable of capturing heteroskedastic phenomena, typical of financial time series data (Maciel & Ballini, 2009). Overall, this research has demonstrated the potential of ANNs as a beneficial tool for individual investors in their search for superior returns on the financial markets.

3.4.2. Recurrent Neural Networks (RNNs)

In this section, we aim to provide an overview of Recurrent Neural Networks (RNNs), which are a type of artificial neural networks specifically developed for identifying patterns within sequential data. To achieve this, we will summarize the key points presented by Lipton (2015) in his publication titled "A Critical Review of Recurrent Neural Networks for Sequence Learning”. The objective here is to grasp the fundamental workings of RNNs, without developing extensive mathematical considerations, as it falls beyond the scope of this master's thesis.

3.4.2.1. Activation functions

In the previous section, in which we discovered the architecture of multilayers perceptrons, we have seen that each neuron j from any layer l has a specific activation value denoted a^L resulting from the application of the activation function σ, which is sometimes called the link function. However, this σ is a widely used specific activation function for multilayer perceptrons known as sigmoid.

$$\sigma(\mathrm{z}) = 1/(1 + \mathrm{e}^{-\mathrm{z}})$$

To understand the functioning of Recurrent Neural Networks, it is important to consider other activations functions. Common choices for the activation function of RNNs also include the sigmoid, but also and the $\tanh$ function, that can be found in various RNN models:

$$\phi(\mathrm{z}) = (\mathrm{e}^{\mathrm{z}} - \mathrm{e}^{-\mathrm{z}})/(\mathrm{e}^{\mathrm{z}} + \mathrm{e}^{-\mathrm{z}})$$

Another activation which has become widely used for developing Recurrent Neural Networks is the rectified linear unit (ReLU) function.

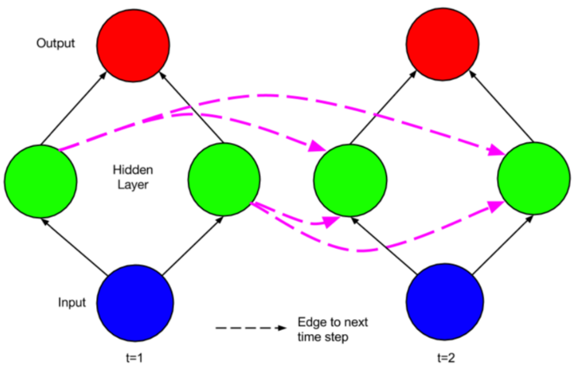
$$\mathrm{ReLU_j(z)} = \max(0, \mathrm{z})$$

It should also be noted that it is common standard to use a specific activation function for the last layer (output layer). A popular function in that regard is the softmax function. However, this function is designed for multi-class classification tasks, which are out of the scope of this master thesis. For stock returns prediction, linear activation (such as the identity function) could constitute a good choice as the activation function of the output layer.

3.4.2.2. Recurrent Neural Networks specificities

Unlike feedforward neural networks like Multilayer Perceptrons, RNNs maintain an internal state that can allow information to persist from an arbitrarily long sequence of previously seen inputs (Figure 7). These characteristics make this type of machine learning model very suitable for time series analysis and has by nature a strong potential for stock returns predictions.

Figure 7.
Recurrent Neural Networks Theoritical Representation.



Note. Source: Lipton, Z. C. (2015). A critical review of recurrent neural networks for sequence learning.

A simple RNN network, just like the Multilayer Perceptrons model we described above, consists of an input layer, one or plural hidden layer(s), and an output layer. A key distinction between MLPs and RNNs however is that in an RNN, the hidden state at time step t not only depends on the current input $\mathrm{x(t)}$ but also on the previous hidden state $\mathrm{h(t-1)}$. This recurrent connection allows the network to capture temporal dependencies in the data and exhibit dynamic behavior. The hidden state at time t, denoted as $\mathrm{h(t)}$, can be updated using the current input $\mathrm{x(t)}$ and the previous hidden state $\mathrm{h(t-1)}$ according to the following equations. In the presented model, the chosen activation function for the hidden layers is the tanh function, while the activation function for the output layer is the identity function.

$$h^t = \phi(W_{hx} * \mathrm{x(t)} + W_{hh} * h^{t-1} + b_h)$$

$$\hat{\mathrm{y}}^t = \left(W_{yh} * h^t + b_y\right)$$

In the above equations, W_{hx} and W_{hh} are the weight matrices for input-to-hidden and hidden-to-hidden connections respectively, and b_h and b_y are bias terms.

In the previous section about Multilayer Perceptrons, we reviewed in detail the main concepts and fundamental underlying equations of the backpropagation algorithm. A similar approach has been designed to train Recurrent Neural Networks, named Backpropagation Through Time (BPTT). It extends the backpropagation algorithm to handle the temporal nature of RNNs. This allows the RNN to learn from and adjust its predictions based on the entire sequence of input data, rather than making predictions based on a static pre-trained model.

3.4.2.3. Shortcomings of RNNs

Despite their advantages for time series forecasting we have already discussed, RNNs have been known to be difficult to train. The main challenge in training RNNs is the problem of vanishing or exploding gradients, which refers to the decay or blow up of gradients as they are propagated back in time during the BPTT algorithm. This problem makes it hard for the RNN to learn long-range dependencies in the data. Therefore, the standard RNN model we described can only find applications for predictions influenced mostly by short-term dependencies. One of the objectives of this master thesis is to determine empirically whether this characteristic makes simple RNNs a strong or weak candidate for short-term stock returns predictions.

3.4.3. Long Short-Term Memory (LSTM)

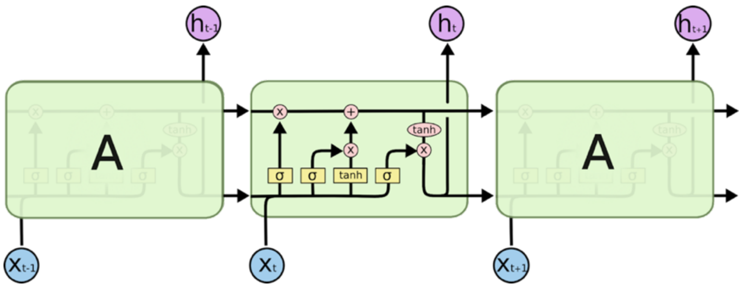
Long Short-Term Memory (LSTM) is an advanced type of RNN developed to address the vanishing gradient problem we just described, allowing thus mode to learn long-term dependencies in data. This section will try to summarize the content of the brilliant blog post written by Olah (2015) on "Understanding LSTM Networks".

3.4.3.1. Structure of Long Short-Term Memory networks

Long Short-Term Memory networks (LSTMs) indeed address the long-term dependency problem and the gradient problem shown by standard RNNs, as discussed above. LSTMs, introduced by Hochreiter & Schmidhuber (1997), have shown remarkable performance on a broad spectrum of tasks, mainly due to their unique ability to retain and manage information over longer periods than standard Recurrent Neutral Networks.

LSTMs are composed of a chain of neural network modules, each containing four interacting layers that collectively enable the learning of long-term dependencies. The key feature of LSTMs is their 'cell state', a vector for information that flows through the network, influenced by subsequent 'gates'. These gates regulate the flow of information by deciding what information to keep or discard, thereby ensuring a well-adjusted balance between the old and the new data (Figure 8).

Figure 8.
Long Short-Term Memory Networks Theoritical Represention.



Note. Source: Olah, C. (2015, August 27). *Understanding LSTM Networks*.

The 'cell state' is a horizontal line running through the LSTM structure, often regarded as a *'conveyor belt'* for information. It carries data throughout the network with minimal changes, ensuring that relevant information can be accessed over long sequences. 'Gates', which include the forget gate, input gate, and output gate, are the control mechanisms for the LSTM. They decide what information is retained, discarded, or generated at each time step, using Sigmoid or Tanh activation functions.

The 'forget gate' determines the amount of past information to forget, while the 'input gate' decides which new information from the current input should be stored. These gates interact with the cell state, effectively mitigating the vanishing gradient problem by preserving the information over long periods, enabling the LSTM to learn long-term dependencies. Finally, the 'output gate' filters the cell state information to decide the final output, thus regulating the data flowing out of the network at each time step. These elements collectively give LSTMs their superior performance in handling sequential data with long-term dependencies, which is particularly helpful for stock returns prediction.

3.4.3.2. Application of LSTM networks for financial market predictions.

The literature about the role and performance of LSTM to predict financial time series is vast. In their paper, Fischer and Krauss (2018) apply Long Short-Term Memory (LSTM) networks to predict financial market performance, specifically focusing on the S&P 500 from December 1992 to October 2015. The authors make three key contributions:

In that study, the authors found that the LSTM model outperforms standard deep networks, logistic regression, and most of the time, the random forest model. They also investigated the "black box" nature of LSTMs, finding common patterns of stocks selected for profitable trading. These stocks typically exhibited below-mean momentum, strong short-term reversal characteristics, high volatility, and high beta – all factors that are related to existing capital market anomalies (Fischer & Krauss, 2018). Finally, they implemented a simplified rules-based trading strategy based on the common patterns found in the LSTM portfolio. This strategy achieved about 50% of the LSTM returns, further confirming the effectiveness of LSTM networks at identifying winning patterns.

The authors conclude that LSTM networks effectively extract meaningful information from noisy financial time series data, outperforming other models in terms of prediction accuracy and daily returns after transaction costs. Therefore, they state that deep learning, particularly in the form of LSTM networks, is a significant advancement in financial markets prediction.

3.4. Detailed methodology for optimal portfolio construction and portfolio analysis

The methodology chosen for solving the mean-variance optimization problem can be broken down into several key stages. Some of the procedures described are repeated for all the considered regression models, in particular the data collection and preparation, as well as the portfolio optimization and performance analysis compartments. The model development and implementation section will however be specific to each model considered in the context of this master thesis:

- Linear regression model
- Ridge regression model
- Lasso regression model
- Elastic net regression model
- Multilayer perceptrons regression model
- Recurrent neural networks (RNN) regression model
- Long Short-Term Memory regression model

3.4.1. Data collection and preparation

The first stage involves data collection and preparation. Macroeconomic data is gathered from the Federal Reserve Economic Data (FRED) database using the FRED API, with an API key provided for this purpose. A selection of economic indicators is chosen based on their relevance given the chosen investment universe, namely the Dow Jones Industrial Average Index constituents. These indicators encompass a broad array of macroeconomic indicators which are all described in section {4.2.3.}.

The collected data spans from January 1, 2001 to January 1, 2021, ensuring a comprehensive overview of economic trends over the last two decades. As we are working with monthly observations, it is crucial to work with a considerable time window to construct a large enough sample to train our various models. The data is then cleaned, and missing values are filled in, where possible, using the Last Observation Carried Forward method (LOCF). Finally, the data is consolidated into a single data frame for easier manipulation in later stages.

Additionally, stock price data for all components of the Dow Jones Industrial Average (DJIA) is extracted from an Excel spreadsheet which has been constructed with help of the Refinitiv Eikon Excel Add-in. The time window considered should be large enough to enable merger with the macroeconomic dataset constructed earlier. This data is then filtered to match the same date range as the macroeconomic data that was extracted from the FRED database.

Then, the macroeconomic and financial datasets are merged. The macroeconomic data and the DJIA component data are joined together, creating a single, unified dataset. This comprehensive dataset is split into two separate data frames for further analysis: one for the stock prices and another for the macroeconomic covariates. The log returns for the stock prices are then calculated, and the comprehensive dataset is updated. Log returns, as opposed to simple returns, are preferred due to their time-additive property, making them particularly suited for the subsequent time series analysis.

3.4.2. Model development and implementation

As stated earlier, this step will have a different instance for each regression model considered in the context of this master thesis. For each regression type, and for each DJIA component, a new regression model is trained and a prediction is made based on an out-of-sample approach to predict one-step ahead returns for the last 36 periods in our dataset. This results in a robust set of predictions that take into account the unique behaviors and characteristics of each individual stock.

The models are trained using a rolling window approach, meaning that the model training is continuously updated as new data becomes available. Each rolling window encompasses a certain number of previous periods, taking into account both the most recent data and a sequence of preceding data points. This method allows each of our models to effectively capture both short-term and long-term temporal dynamics in the data. The output, for each model, is thus a list of predicted returns for each stock in the DJIA over the last 36 periods.

3.4.3. Portfolio optimization

The next stage of the methodology involves portfolio optimization using the mean-variance optimization framework as suggest by Markovitz (1952). The predicted returns that we estimated at the previous step are used as input into this optimization model. Portfolio weights are determined using a solver that determines the optimal portfolio weights maximizing the Sharpe ratio, which is a measure of the risk-adjusted return. These weights thus represent the proportion of the total investment that should be allocated to each DJIA stock to reach the highest achievable Sharpe ratio.

Given that financial markets can be highly volatile and that regression models predictions are continuously updated, the portfolio's composition may change over time to adapt to new market conditions and maintain an optimal risk-return trade-off. This dynamic approach aligns well with the unpredictable nature of financial markets, providing a robust strategy.

3.4.4. Strategy Performance Analysis

In this final section, the performance of each of the seven regression models-based strategies is evaluated and compared. This involves calculating the realized return for each model in the 36 last time steps of our dataset, for each of the components of the Dow Jones Industrial Average (DJIA) index. The realized return is computed by multiplying the model-optimized portfolio weights by the actual return of each stock in the subsequent period. The same procedure is applied to a simple equal-weighted (EW) portfolio for comparison purposes.

The calculated realized returns for each strategy are then compounded over time to produce cumulative returns, which are displayed in a comparative plot. This visualization provides an overview of how each strategy would have performed over the evaluation period. To deepen the performance analysis, several key risk and return metrics are calculated for each strategy. These include the Sharpe ratio, which measures risk-adjusted performance; the Maximum Drawdown, representing the largest drop from peak to bottom; the Sortino ratio, similar to the Sharpe ratio but penalizing downside volatility; and the Compound Annual Growth Rate (CAGR), providing the mean annual growth rate of an investment over a specified period of time. The computation of these metrics contributes to a more comprehensive view of each strategy's performance, taking into account both returns and risks.

3.5. Detailed methodology for the ESG-optimized portfolio construction

3.5.1. Global procedure description

The procedure to integrate ESG considerations in our portfolio construction process follows the results obtained in the previous section, as the generated predicted returns are needed in this analysis. For the sake of simplicity and clarity, it has been decided to only focus on the impact of integrating ESG preferences of investors in the context of the strategy implied by the use of the multilayer perceptrons regression model to predict stock returns.

In this section, we will perform a case study on how to integrate ESG considerations in portfolio construction at a specific date in the past. Specifically, the investment decision that we will investigate could have taken place on November 30, 2020, which corresponds to the penultimate data row of our comprehensive dataset. Considering the penultimate data row allows to go even deeper in our analysis by performing an ex-post analysis and see what are the realized consequences of integrating ESG considerations in portfolio construction.

3.5.2. Integration of ESG considerations from the optimal mean-variance portfolio

We start with the same investment universe as in the previous section: the constituting stocks of the Dow Jones Industrial Average (DJIA). For each asset, we append an ESG score, a composite measure reflecting the company's performance in environmental, social, and governance matters. For this master thesis, the Refinitiv ESG scores have been chosen for the reasons detailed in section 3.2.5. The scores are collected into a dedicated vector from the Refinitiv Eikon database using the Excel Add-In of the application.

While the approach described further is robust, it is important to acknowledge one key limitation: the use of static ESG scores. The choice to use static ESG scores was made for simplicity reasons as these scores, by their nature, do not exhibit significant changes over time compared to the dynamic nature of other metrics we employ, particularly stock returns and their volatility. This nuance however highlights the need for careful interpretation and consideration of the static ESG scores within the broader context of our analysis.

First, we start by constructing the mean-variance-ESG efficient surface as described by Steuer and Utz (2023). This three-dimensional object helps the investor make an informed investment decision by offering an intuitive visualization of the trade-off he/she has to make in terms of risk, return, and portfolio ESG impact. In a first time, we will not apply any ESG constraint on the computation of the efficient surface.

Then, we will attempt to understand the features of the optimal mean-variance portfolio, in particular its Sharpe ratio and ESG aggregate score. This will be the starting point of our ESG integration procedure. From there, we will attempt to understand how much return an investor has to give up in order to achieve a desired level of ESG integration. This ESG integration will be measured with help of the ESG Integration Index proposed by Steuer and Utz (2023), and thoroughly described in section 2.1.5. of the literature review. This index aims at characterizing the ESG score improvement of a portfolio, by placing it on a scale that ranges from the mean-variance optimal portfolio (ESG Integration Index = 0) to the portfolio that provides the highest attainable ESG score (ESG Integration Index = 100).

We can already comment at this stage that our approach is the inverse of the one described by Steuer and Utz (2023): in their study, the authors try to characterize the achievable ESG score improvement based on the number of basis points that the investor is willing to give up to obtain a greener portfolio. On the other hand, our approach is an attempt to understand, for a desired level of ESG integration, how much a serious ESG investor needs to give up both at portfolio allocation stage but also in terms of ex-post realized returns. One of the research objectives of this master thesis is to check whether the same conclusions can be drawn by using this different but coherent approach.

4. Development and empirical results

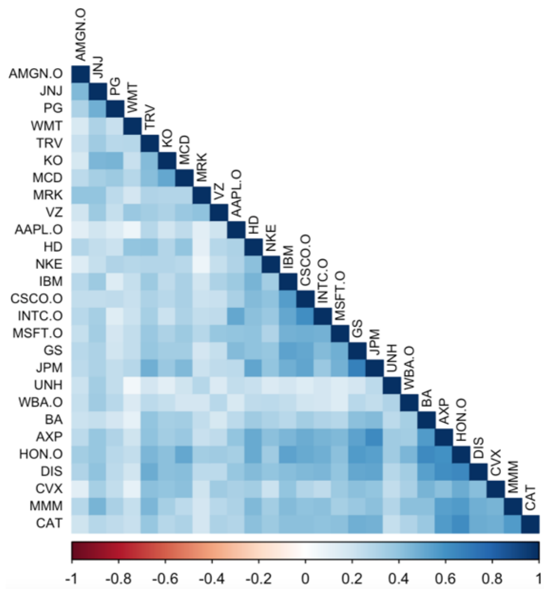
4.1. Financial data description

In order to provide context for our empirical analysis, it is important to describe our stocks dataset. In this section, we will examine the relationships between the returns of the assets within our investment universe. We will visualize this relationship through a variance-covariance matrix, which will help us map the positioning of each asset in terms of return and risk. Additionally, we will evaluate the risk and return levels and Sharpe ratio of each stock to gain insights about the risk-return tradeoff they present. By considering these factors, we can better understand the characteristics and performance of our stocks before diving into the results of our empirical analysis.

4.1.1. Variance-covariance matrix

The variance-covariance matrix is a square matrix that summarizes the variances and covariances between the individual stocks in the portfolio (Figure 9). It provides valuable information about the risk and diversification potential of the portfolio. The diagonal elements of the matrix represent the variances of the individual stocks, indicating their individual risk levels. Any off-diagonal element represents the covariances between a combination of two stocks, reflecting the extent to which their returns move together. By analyzing the variance-covariance matrix, we can assess the overall risk of the portfolio and make informed decisions about asset allocation and diversification to optimize their risk-return trade-off.

Figure 9.
Variance-Covariance Matrix of the Defined Investment Universe.



Note. Source: Refinitiv Eikon, FRED

The variance-covariance matrix is a fundamental component of our methodology, particularly in the portfolio optimization process. It serves as the primary measure of portfolio risk, allowing us to calculate the Sharpe ratios for various portfolios within our predefined investment universe. By utilizing the matrix, we can identify the optimal portfolio with the highest Sharpe ratio.

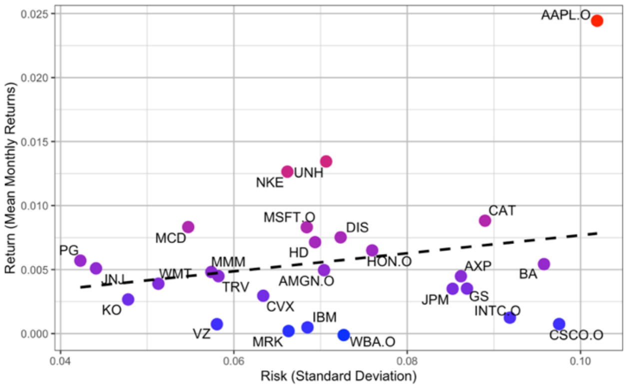
From the provided information, it is evident that there is a positive correlation among the returns of all stocks, with higher correlation observed among stocks within the same industry sector.

This understanding of correlations helps our portfolio optimization algorithm to make informed decisions about portfolio construction and diversification.

4.1.2. Risk/return profiles of the considered stocks

Figure 10 compares the average historical levels of return and risk for each asset within our investment universe over a 20-year time window. The line running through the plot represents the coefficient of a linear regression between the vector of risks and the corresponding vector of returns, with a slope of 0.070528.

Figure 10.
Risk/Return Trade-Off of the Investment Universe Constituents.

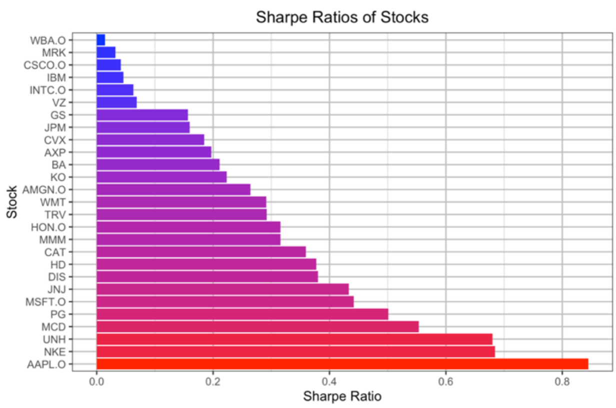


Note. Source: Refinitiv Eikon, FRED

4.1.3. Sharpe ratios of the considered stocks

Figure 11 illustrates the Sharpe ratios of the constituents within our investment universe. It is worth noting that a risk-free rate of -0.05% was used for these calculations, based on the ECB interest rate at the end of the selected prediction window (2018-2021). From the graph, it can be inferred that Apple Inc. (AAPL.O) offers the most favorable return-to-risk tradeoff, indicating superior risk-adjusted performance compared to other constituents. On the other hand, Walgreens Boots Alliance, Inc. (WBA.O) demonstrates the weakest performance, as it barely achieves positive Sharpe ratio despite the consideration of a negative interest rate. These observations highlight the differing risk-return characteristics and relative performance of the stocks within our investment universe.

Figure 11.
Sharpe Ratios of the Investment Universe Constituents.



Note. Source: Refinitiv Eikon, FRED

4.2. Investment strategies performance analysis

In this section, we aim at making an in-depth comparison of the out-of-sample returns implied by the seven selected investment strategies, and ultimately to identify the one which stands out. As a reminder, each of these strategies rely on a different regression method to predict future stock returns, impacting the asset allocation suggested by the mean-variance optimization problem:

- Linear regression
- Ridge regression
- Lasso regression
- Elastic net regression
- Multilayer perceptrons (MLP) regression
- Recurrent Neural Networks (RNN) regression
- Long Short-Term Memory (LSTM) neural networks regression

4.2.1. Cumulative returns

Cumulative returns refer to the total return or gain for a given investment over a specified period. It represents the aggregate performance of the investment, considering the compounding effect of returns over time. For this master thesis, as explained earlier, it has been chosen to compare the selected investment strategies using a 3-years wide out-of-sample prediction approach.

Figure 12.
Comparison of Cumulative Returns for the different Investment Strategies.

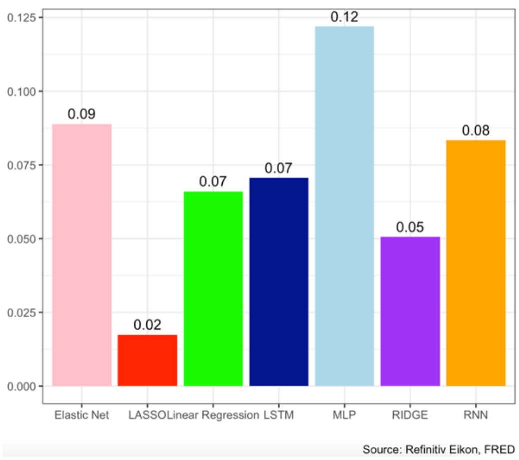
Source: Refinitiv Eikon, FRED

Figure 12 clearly demonstrates that the Multilayer Perceptron (MLP) strategy outperforms the other strategies in terms of cumulative returns. The MLP strategy shows the highest cumulative returns at the end of the considered window, indicating its effectiveness in capturing investment opportunities and delivering consistent gains over time. Following MLP, the elastic net model implied strategy demonstrates a relatively strong performance. Linear Regression, LSTM, Ridge, and RNN implied strategies show varying degrees of cumulative returns. However, it is important to note that the lasso implied strategy is the only one that underperformed the buy-and-hold (equal weight) strategy. This underperformance can be explained partially by a poor choice of asset allocation right before the 2020 sanitary crisis crash. The results of this section highlight the potential advantages of employing the MLP strategy for investment decision-making.

4.2.2. Compound annual growth rate

The compound annual growth rate, often abbreviated as CAGR, is a measure used to calculate the average annual growth rate of an investment over a specified time period, assuming that the investment grows at a constant rate each year. CAGR accounts for compounding effect, meaning that this metrics considers the reinvestment of returns to generate higher growth. This characteristic makes the CAGR a good candidate for performance comparisons across diverse investment strategies.

Figure 13.
Strategies Comparison based on CAGR.



The investment strategies examined in the study exhibited different Compound Annual Growth Rates (CAGR) as we can see on Figure 13, and we can see that their rank in terms of this metric could also have been deducted by reading the cumulative returns graph (Figure 12). The RNN strategy achieved a CAGR of 8%, while Linear Regression and LSTM strategies both recorded a CAGR of 7%. The MLP strategy demonstrated the highest CAGR at 12%, followed by the elastic net strategy with 9%. The ridge strategy attained a CAGR of 5%, while the lasso strategy exhibited the lowest CAGR at 2%.

4.2.3. Maximum Drawdown

Maximum Drawdown refers to the largest decline or loss experienced by an investment or portfolio during a specific period, measured from its highest value to the lowest value before a recovery takes place. It represents the maximum percentage drop in value from a previous peak to the subsequent trough. This metric helps investors assess the magnitude of the worst decline suffered by an investment before it starts to regain value. Maximum Drawdown is a common risk measure used to evaluate the potential downside and volatility of an investment strategy or asset.

Figure 14.
Strategies Comparison based on Maximum Drawdown.

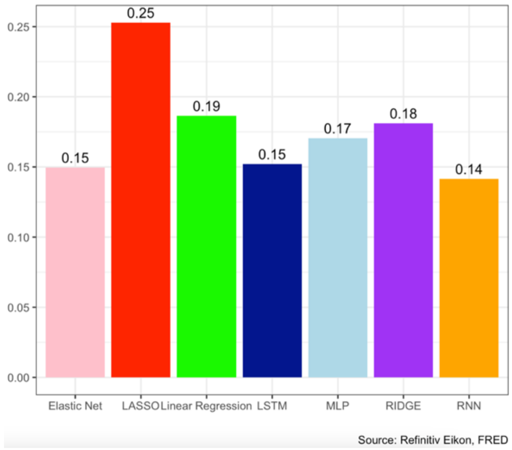
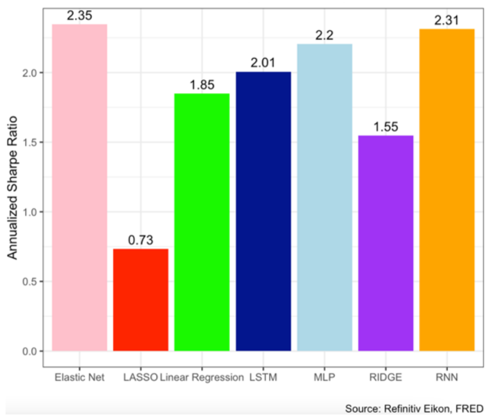


Figure 14 shows that lasso strategy yields the maximum drawdown of 25%, indicating a relatively higher level of downside risk. Linear regression and ridge strategies exhibit maximum drawdowns of 19% and 18% respectively, implying a slightly lower degree of volatility. Moving towards the MLP and LSTM strategies, their maximum drawdowns stand at 17% and 15% respectively, suggesting a comparatively reduced level of risk. Both the Elastic Net and RNN strategies demonstrate the lowest maximum drawdowns, with values of 15% and 14% respectively, implying a higher level of stability and lower downside risk. Overall, while lasso presents the highest risk in terms of maximum drawdown, RNN stands out as the strategy with the lowest maximum drawdown, offering potentially more stable returns, and potentially demonstrating as the less vulnerable strategy in case of extreme events.

4.2.4. Sharpe ratio

The Sharpe ratio is a measure used to evaluate the risk-adjusted return by quantifying the excess return generated by an investment in relation to its volatility or risk. It is calculated by subtracting the risk-free rate of return from the investment's average return and dividing the result by the investment's standard deviation. The Sharpe ratio allows investors to assess whether the returns achieved by an investment compensate for the level of risk undertaken. A higher Sharpe ratio indicates a better risk-adjusted performance, suggesting that the investment has provided higher returns relative to its level of risk.

Figure 15.
Strategies Comparison based on Annualized Sharpe Ratio.

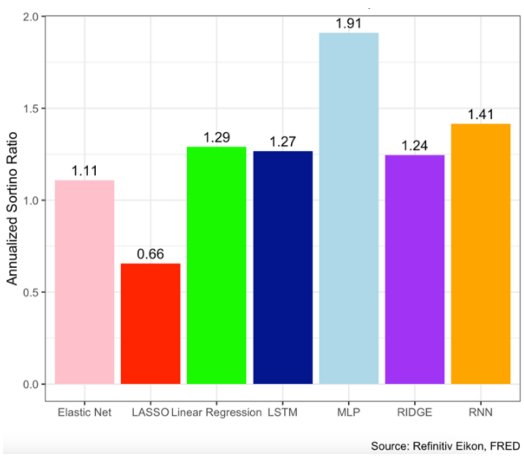


On Figure 15, it can be observed that among the considered strategies, the elastic net technique demonstrates the highest Sharpe ratio of 2.35, indicating a favorable risk-adjusted return. Following closely is the recurrent neural network (RNN) approach, with a Sharpe ratio of 2.31, indicating a strong risk-adjusted performance as well. The multilayer perceptron (MLP) strategy also exhibits a respectable Sharpe ratio of 2.2. On the other hand, the long short-term memory (LSTM) model achieves a Sharpe ratio of 2.01, indicating a slightly lower risk-adjusted return compared to the previous strategies. Linear regression presents a Sharpe ratio of 1.85, while the ridge regression approach achieves a Sharpe ratio of 1.55, reflecting a lower risk-adjusted return compared to previous strategies. Finally, the lasso method exhibits the lowest Sharpe ratio among the strategies analyzed, at 0.73, indicating a poor risk-adjusted performance as a Sharpe ratio smaller than 1 is often considered as bad.

4.2.5. Sortino ratio

The Sortino ratio is also a risk-adjusted performance measure that focuses on downside risk, specifically the level of downside deviation. The ratio measures the excess return of an investment or portfolio relative to its downside deviation, taking into account only the negative returns. Similar to the Sharpe ratio, a higher Sortino ratio indicates a better risk-adjusted performance. It is calculated by subtracting the risk-free rate of return from the investment's average return and dividing the result by the investment's downside deviation. The Sortino ratio is particularly useful for investors that exhibit high risk aversion and who are more concerned about minimizing downside risk and preserving capital rather than overall volatility. It provides a more focused assessment of an investment's performance in unfavorable market conditions.

Figure 16.
Strategies Comparison based on Sortino Ratio.



Among the strategies examined on Figure 16, the MLP (Multilayer Perceptron) approach presents the highest Sortino ratio of 1.91, indicating a strong risk-adjusted return. The RNN (Recurrent Neural Network) method follows closely with a Sortino ratio of 1.41. The LSTM (Long Short-Term Memory) model achieves a Sortino ratio of 1.27. Both linear regression and ridge regression techniques exhibit similar Sortino ratios of 1.29 and 1.24, respectively, showcasing still decent risk-adjusted performances. The elastic net technique demonstrates a Sortino ratio of 1.11. Finally, the lasso method exhibits the lowest Sortino ratio of 0.66, indicating a relatively poorer risk-adjusted performance as compared to the other strategies analyzed.

4.2.6. Treynor ratio

The Treynor ratio is another financial performance measure used to evaluate the risk-adjusted return of an investment or portfolio by accounting for systematic risk, often measures as beta, the coefficient of the market returns regressor in a linear regression with the returns of the portfolio under consideration as the dependent variable. It quantifies the portfolio return generated per unit of systematic risk undertaken. The ratio is calculated by subtracting the risk-free rate of return from the investment return and dividing the result by the beta of the investment. A higher Treynor ratio suggests a better risk-adjusted performance, indicating that the investment has generated more excess return for each unit of systematic risk. It is commonly used to assess the performance of investment portfolios in relation to their exposure to market risk.

Figure 17.
Strategies Comparison based on Treynor Ratio.

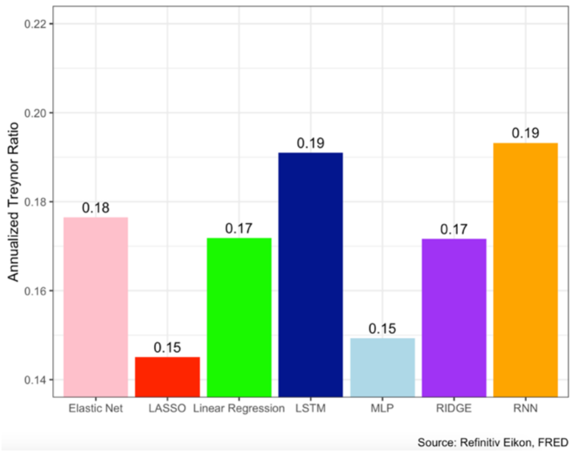
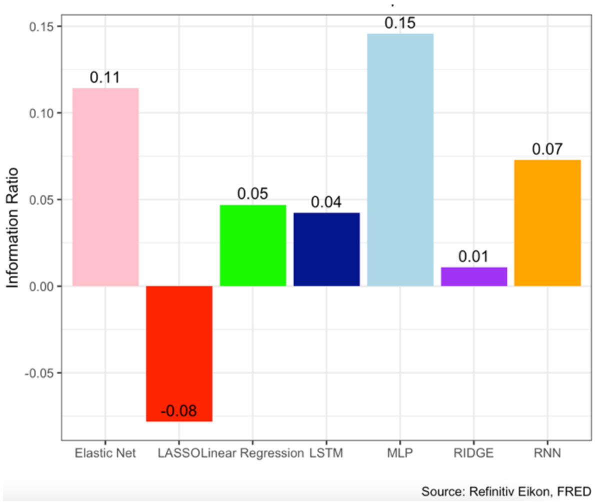


Figure 17 evaluates the performance of different investment strategies by analyzing their Treynor ratio values. Among the strategies examined, the RNN and LSTM methods achieve the highest Treynor ratio of 0.19, indicating the best risk-adjusted returns in the strategies sample. The elastic net technique, linear regression and ridge regression techniques exhibit Treynor ratios of 0.18 and 0.17 respectively. The MLP and lasso methods exhibit a Treynor ratio of 0.15, suggesting relatively poorer risk-adjusted performances compared to the other strategies analyzed.

4.2.7. Information ratio

The information ratio is also a risk-adjusted performance measure, used to evaluate the excess return generated by an investment or portfolio relative to the level of risk undertaken. It compares the investment return over a benchmark or risk-free rate, divided by the volatility of that excess return. A higher information ratio indicates better risk-adjusted performance, implying that the investment manager has generated more excess return per unit of risk. It is commonly used to assess the performance of actively managed portfolios and strategies, which particularly relevant in the context of this study. The benchmark used to compute the excess returns is the portfolio implied by the buy-and-hold strategy.

Figure 18.
Strategies comparison based on Information Ratio.



From Figure 18, it can be understood that the MLP method demonstrates the highest Information ratio of 0.15, indicating that it is the strategy with the strongest risk-adjusted return according to this metric. The elastic net technique follows with an Information ratio of 0.11, followed by the RNN method which exhibits an Information ratio of 0.07. The linear regression and LSTM approach achieve respectively Information ratio of 0.05 and 0.04. The ridge regression technique demonstrates a marginal Information ratio of 0.01, reflecting a minimal risk-adjusted performance. In should be noted that the lasso method exhibits a negative Information ratio of -0.08, which is not surprising as it underperformed the benchmark buy-and-hold strategy.

4.2.8. Performance analysis synthesis

The Multilayer Perceptron (MLP) strategy consistently stood out as a robust performer across multiple measures, making it an informed choice for the succeeding section. With the highest cumulative returns and Compound Annual Growth Rate (CAGR) of 12%, along with the top Sortino ratio of 1.91 and Information ratio of 0.15, the MLP strategy demonstrated remarkable efficiency in generating returns, compensating for risk taken, and delivering risk-adjusted performance relative to the buy-and-hold benchmark.

The Elastic Net Regression model also presented notable performance, achieving the second-best CAGR of 9% and the highest Sharpe ratio of 2.35, reflecting a favorable risk-adjusted return. This strategy additionally exhibited lower downside risk with a maximum drawdown of 0.15.

Meanwhile, the Recurrent Neural Networks (RNN) and Long Short-Term Memory Neural (LSTM) Networks strategies displayed balanced performance across the board. Both techniques shone in terms of their Treynor ratios of 0.19, indicating efficient risk-adjusted returns considering their systematic risk (beta).

The Linear Regression method exhibited mixed results, with a CAGR of 7% and respectable risk-adjusted performance metrics, including a Sharpe ratio of 1.85, a Sortino ratio of 1.29, and a Treynor ratio of 0.17.

In contrast, the Ridge and Lasso Regression strategies demonstrated weaker performances compared to the other methods. Ridge Regression registered modest returns with a CAGR of 5%, alongside average risk-adjusted metrics. Lasso Regression, however, was behind all the strategies in terms of return generation and risk management. With the lowest CAGR of 2%, the highest maximum drawdown of 0.25, and the smallest Sharpe, Sortino, Treynor, and Information ratios, Lasso Regression showed the greatest downside risk and poorest risk-adjusted returns.

In a nutshell, while most strategies hold their unique strengths and weaknesses, the MLP strategy seems to offer the most compelling balance of strong return generation and efficient risk management. Consequently, we will be proceeding with the MLP strategy for further investigation and implementation in the next section.

4.3. Inclusion of ESG preferences

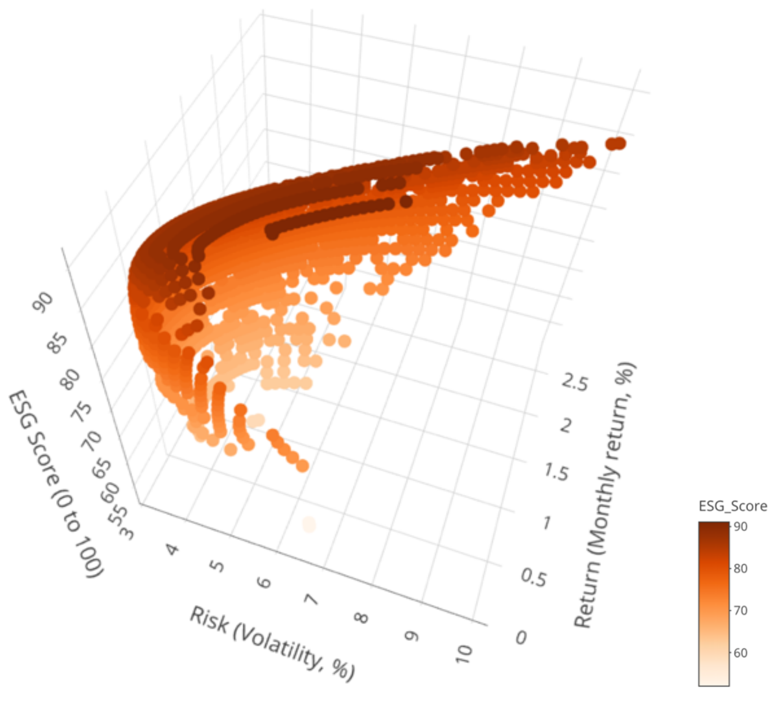
In this section, we study portfolio optimization beyond the traditional mean-variance optimization problem, by exploring an alternative procedure to account for ESG considerations in the portfolio optimization process.

4.3.1. The Mean-Variance-ESG Efficient Surface

Our analysis begins by establishing a foundation for subsequent discussion, by outlining the framework that will guide our exploration. This framework, mainly inspired by the work of Steuer et al. (2023), adopts the concept of an efficient surface extending in a three-dimensional space, where the axes represent risk, return, and ESG score. The efficient surface visualizes the portfolio with the maximum attainable ESG score for any combined level of risk and return.

The efficient surface displayed below in figure 19 represents the trade-off that a rational investor should consider in his/her asset allocation decisions with regard of our defined investment universe, on November 30, 2020. At this stage, ESG constraints are not accounted for in the efficient surface as the objective is to make an initial ESG-naive investment decision, that will be used as a starting point further in our study.

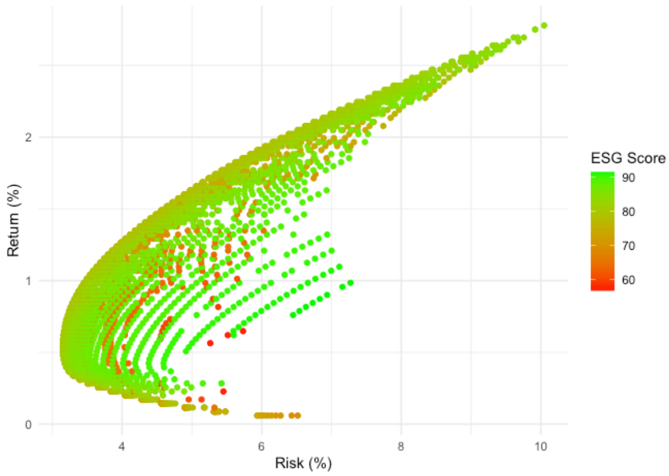
Figure 19.
Mean-Variance-ESG Efficient Surface (Steuer and Utz, 2023).



Note. Source: Refinitiv Eikon, FRED

For an investor more familiar with the traditional mean-variance efficient frontier concept, and eager to draw parallels with it, the Mean-Variance-ESG efficient surface can be visualized as shown below using a color scale (Figure 20). Nevertheless, due to the density of portfolios along the mean-variance efficient frontier, this representation is less appealing than the efficient surface. Consequently, we will not be utilizing this visual representation in our ongoing analysis, and rather prefer the efficient surface representation.

Figure 20.
M-V-ESG Efficient Surface, 2D Representation (Steuer & Utz, 2023).



Note. Source: Refinitiv Eikon, FRED

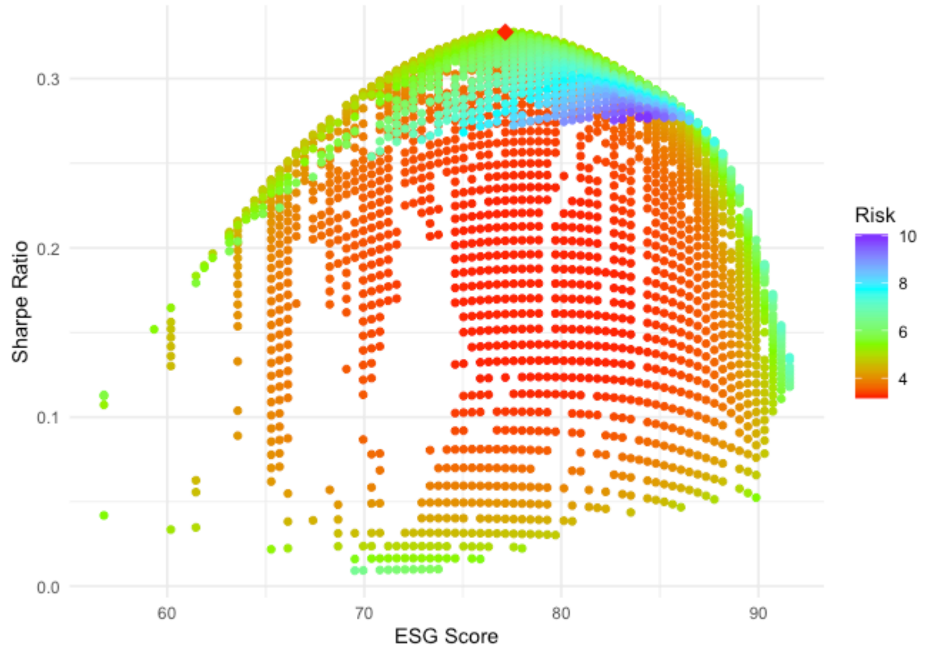
4.3.2. ESG-naive investment decision-making

On Figure 21, we show the ESG-efficient frontier as proposed by Pedersen et al. (2021), enhanced by a color scale indicating a measure of risk for any portfolio in the ESG-efficient frontier to mitigate its primary shortcoming identified by Steuer and Utz (2023): the unique focus on the problem aiming at maximizing Sharpe ratio and subject to ESG constraints, while overlooking the investor's risk profile. With this additional specification, the investor can make a more informed investment decision that is more coherent with its risk profile.

The red dot on Figure 21 shows the portfolio that a non-serious ESG investor would select. It is the optimal portfolio based on the mean-variance model, providing the highest achievable Sharpe ratio. The maximal attainable Sharpe ratio, surprisingly, stands at 0.32 on a monthly scale, resulting in an annualized Sharpe ratio of 1.14. Earlier in section 4.2., we noted that the Sharpe ratio for the MLP model's implied strategy under study was 2.2. This surprising mismatch may push the reader to question the consistency of our results.

Various elements contribute to this apparent discrepancy. During this phase of the ESG integration analysis, the efficient surface and ESG-efficient frontier are constructed based on the predicted returns provided by our MLP regression model. However, when analyzing the performance of different strategies, we used the realized returns implied by each investment strategy to calculate the respective performance metrics. Furthermore, in the evaluation of strategies suggested by machine learning models, we utilized a three-year window of monthly actual returns to compute the Sharpe ratio. In contrast, for the present ESG integration analysis, we're considering only a one-step ahead asset allocation.

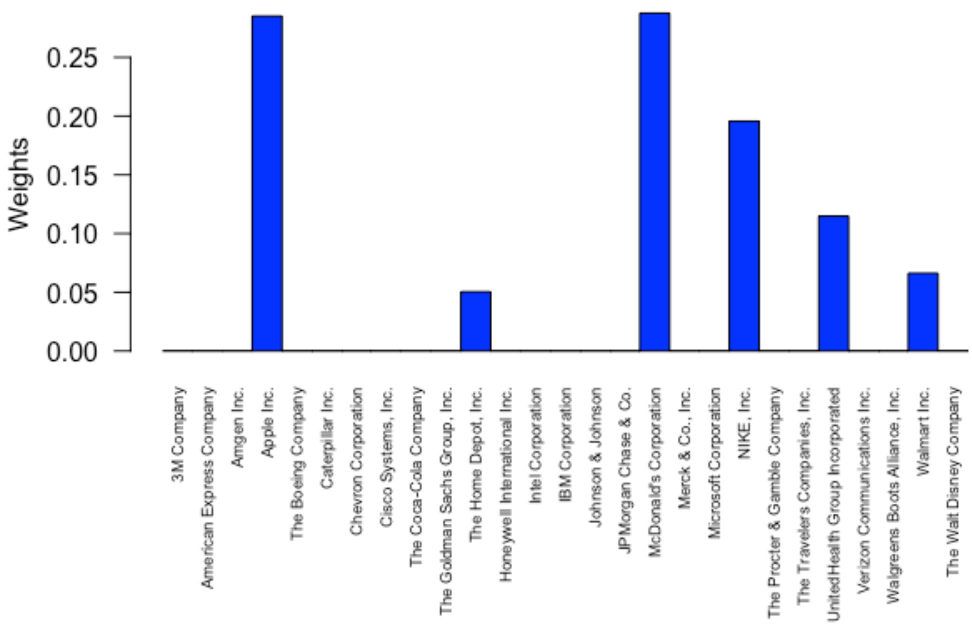
Figure 21.
ESG-Efficient Frontier (Pedersen et al. 2021) with Color Scale to account for Risk Level.



Note: the Sharpe ratio scale is expressed at a monthly frequency.
Source: Refinitiv Eikon, FRED

In the context of this specific investment decision, the optimal portfolio, represented by the red dot in Figure 21, is structured with stock weights as displayed in Figure 22.

Figure 22.
Portfolio Optimal Weights, ESG Inegration Index = 0.



Note. Source: Refinitiv Eikon, FRED

4.3.3. ESG Integration Index and chosen approach

By visualizing the ESG-efficient frontier (Pedersen et al., 2021), we could identify and describe the most preferred portfolio for a non-serious ESG investor. Now, let's dive into the subject of ESG integration in portfolio management: what if the investor is an ESG-serious investor? As a reminder, Steuer and Utz (2023) describe an ESG-serious investor as "*an investor whose strength of interest in ESG makes that consideration a criterion competitive on the playing field of portfolio selection with risk and return, thus causing the investor's efficient frontier to become an efficient surface*".

To quantify the extent of ESG integration this ESG-dedicated investor might choose, we'll employ the ESG Integration Index as proposed by Steuer and Utz (2023), but with minor adjustments to suit our study. This index calculates the degree of ESG integration using the formula:

$$\Delta \upsilon = \frac{\upsilon^T x^{bp,i} - \upsilon^T x^i}{\nu_{max} - \upsilon^T x^i}$$

where υ^T is the ESG Coefficient Vector and x^i is the optimal portfolio without consideration of the ESG preferences of the investor. $x^{bp,i}$ is a portfolio for which a basis point quantity bp has been given up by the investor to achieve higher ESG performance, and ν_{max} is return of the asset allocation that maximizes portfolio ESG score.

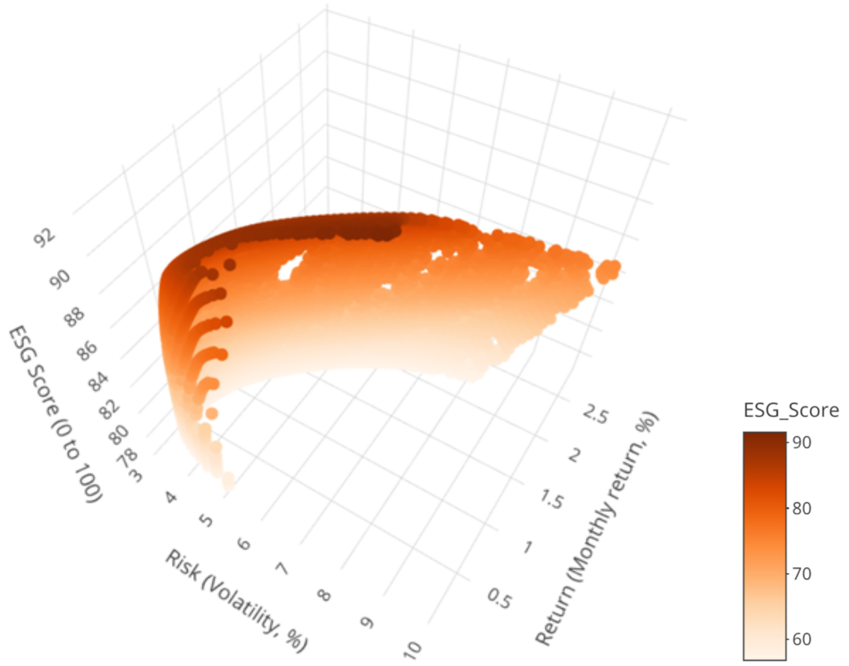
As stated earlier, in this study, we will use a slightly different approach to evaluate the impact of ESG integration in the asset allocation process. Instead of trying to describe the level of ESG Integration improvement that diverse pre-defined levels of expected returns relaxation can yield, we will rather attempt to characterize the needed return sacrifice to achieve a desired level of ESG integration improvement. Thus, the following formula corresponds better to our research objective, which is to describe the evolution of $\upsilon^T x^{bp,i}$ for different values taken by $\Delta \upsilon$.

$$\upsilon^T x^{bp,i} = \upsilon^T x^i + \Delta \upsilon \, \nu_{max} - \Delta \upsilon \, \upsilon^T x^i$$

4.3.4. The trade-off of the ESG serious investor

The optimal mean-variance portfolio, which is the preferred portfolio of a non-serious ESG investor, serves as a starting point for ESG integration. By definition of the efficient surface and the serious ESG investor, both provided by Steuer and Utz (2023), such an investor does not have any incentive to opt for a portfolio with a lower ESG score than the portfolio which maximizes the Sharpe ratio. Thus, the ESG score of the portfolio that maximizes the Sharpe ratio sets the minimum optimal portfolio ESG score. Reading from figure 21, this portfolio exhibits an ESG score of 77, setting the required minimum ESG score at this level (Figure 23).

Figure 23.
M-V-ESG Efficient surface, ESG Integration Index = 0 (Steuer and Utz, 2023).



Note. Source: Refinitiv Eikon, FRED

To make a practical use of the efficient surface representation for ESG investment decision-making, we need a clearer understanding of the investor's ESG preferences. A serious ESG investor might be willing to only drop a few basis points for a greener portfolio in terms of return performance, while others might set ESG integration as an absolute priority, regardless of return cost. To quantify this tendency to integrate ESG considerations, we will use the ESG Integration Index discussed above.

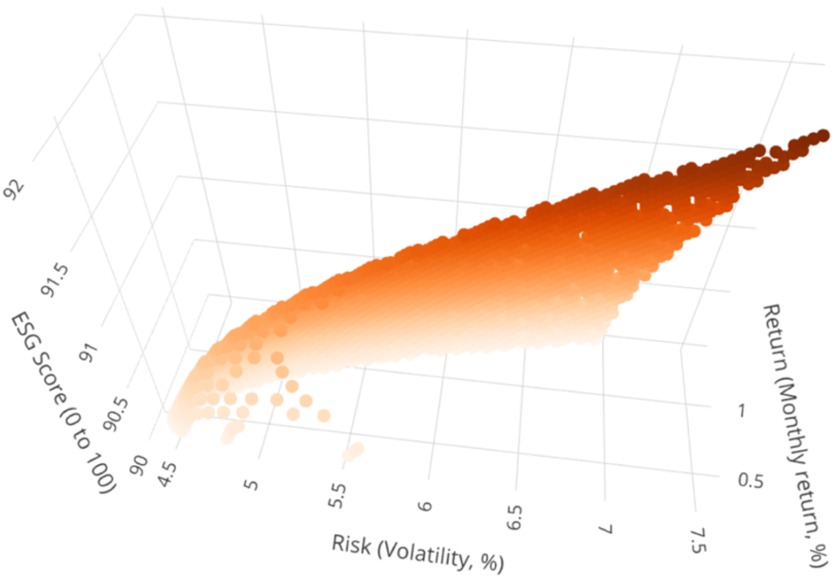
The ESG integration index is defined over a range that goes from 0 to 100. A value of 0 means that the portfolio is characterized by the same ESG score as the optimal mean-variance portfolio, whereas an ESG integration index worth 100 corresponds to the portfolio with the highest attainable Sharpe ratio, which corresponds to a full investment in the stock that has the highest ESG score (which is the company Johnson & Johnson in our study).

Therefore, we will examine various ESG integration levels to comprehend their risk and return implications. Specifically, we will observe the evolution of expected return and risk as the ESG Integration Index increases. Each ESG Integration level is, by definition, associated with a minimum required ESG score.

ESG Integration Index	**0%**	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**	**80%**	**90%**	**100%**
Minimum Required ESG Score	77.1	78.6	80.1	81.6	83.1	84.6	86.1	87.5	89.0	90.5	92.0

For any ESG Integration level, we construct a new efficient surface (Figure 24) that only displays portfolios meeting or exceeding the corresponding minimum required ESG score. The plot below shows the case of the mean-variance-ESG efficient surface for an arbitrarily chosen ESG Integration Index of 87%, with a corresponding minimum required ESG score of 90.

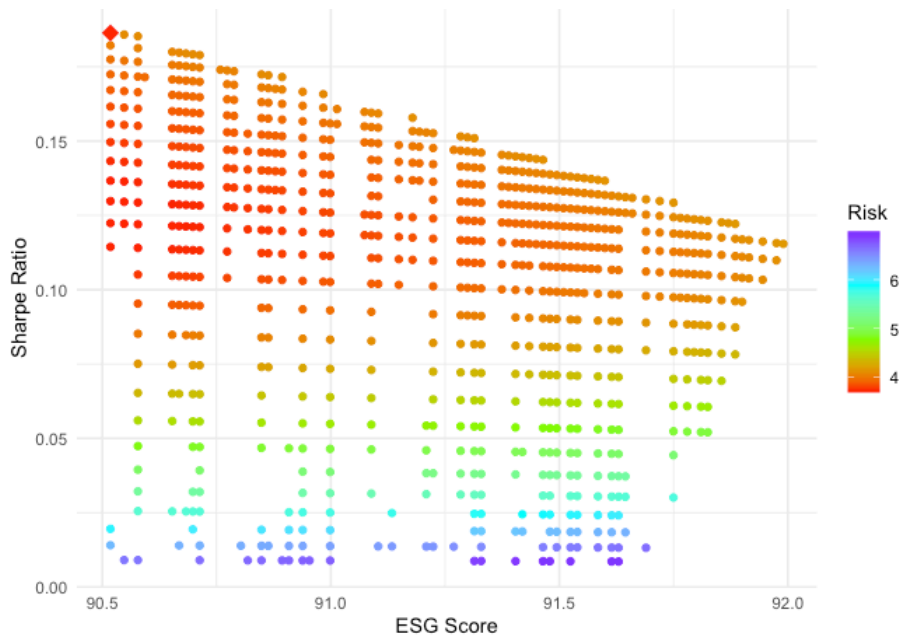
Figure 24.
M-V-ESG Efficient surface, ESG Integration Index = 87% (Steuer and Utz, 2023).



Note. Source: Refinitiv Eikon, FRED

On this plot, it's evident that the efficient surface is considerably limited compared to the initial case, which did not consider any minimal ESG score. Concurrently, it's unsurprising that the corresponding ESG-efficient frontier (Figure 25) advises investing in the portfolio meeting the minimum required ESG score (90.5% in this case), as further increasing the ESG score would significantly reduce the portfolio's Sharpe ratio.

Figure 25.
ESG-Efficient Frontier, ESG Integration Index = 87% (Pedersen et al., 2021).



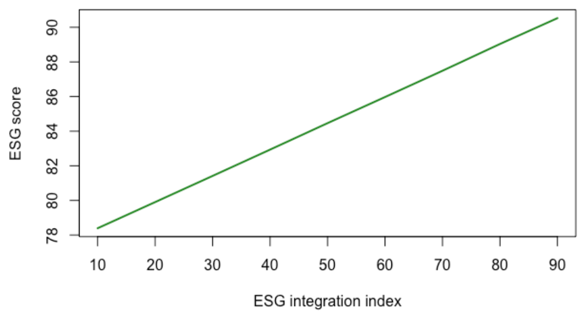
Note: the Sharpe ratio scale is expressed at a monthly frequency.
Source: Source: Refinitiv Eikon, FRED

4.3.5. The cost of ESG integration in asset allocation

Let's now provide a more general approach to understanding the impact of ESG integration, by considering the spectrum of ESG integration levels from 10% to 90%. We will discuss the influence of setting the corresponding minimum required ESG score on returns, risk, and ESG score. The results that will be presented in this sub-section are not specific to a particular investment decision-making instant as we proceeded so far but represent the aggregate results of the last 36 prediction periods from our previously discussed time window.

As expected, on Figure 26, as the ESG Integration Index increases, so does the portfolio's ESG score proportionally. This feature can be attributed to the strictly negative slope of the ESG-efficient frontier after the ESG score that maximized the Sharpe ratio (Figure 21).

Figure 26.
Portfolio ESG Score as a function of ESG Integration Index.

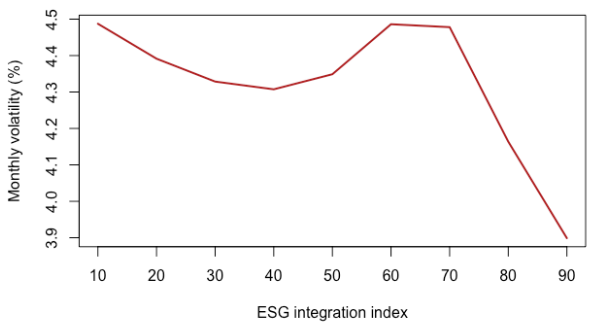


Note. Source: Refinitiv Eikon, FRED

In terms of portfolio risk, the results may seem more surprising. One might assume that as the minimum required portfolio ESG score increases, and with the number of eligible assets decreasing, the diversification opportunities would diminish as the ESG Integration Index rises, leading to an increase in non-systemic risk in our portfolio impacting portfolio expected volatility.

However, as the Figure 27 shows below, selecting the portfolio that maximizes the Sharpe ratio results in a decrease in portfolio risk (measured as expected volatility) for most ESG integration index values. This result may be due to the fact that in our stock sample, the stocks with higher ESG scores typically exhibit lower risk than those with lower ESG scores. Future research could investigate whether this trend persists among a broader group of stocks, or whether this result occurred by chance in this case study.

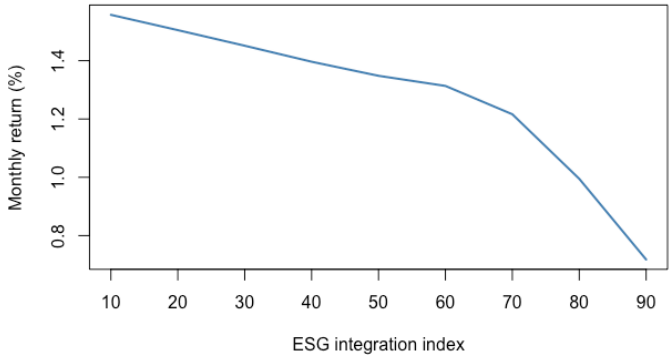
Figure 27.
Portfolio Risk as a function of ESG Integration Index.



Note. Source: Refinitiv Eikon, FRED

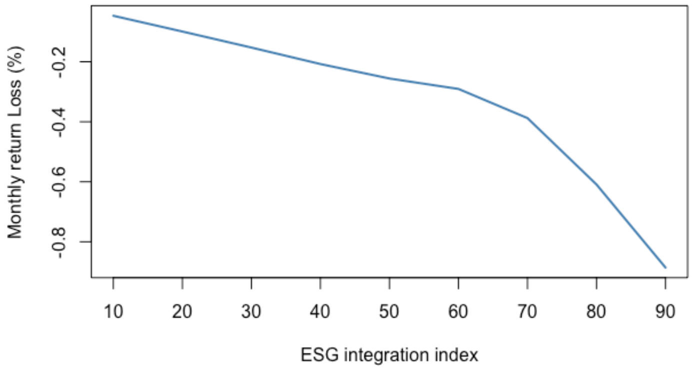
In terms of return performance, we observe a negative impact on portfolio performance with increased ESG Integration level. Figure 28 shows the return implied by the portfolio with the maximum Sharpe ratio for each ESG Integration Index level. Figure 29 and Figure 30 show the same phenomenon, but with the y-axis showing the absolute return loss compared to the mean-variance optimal portfolio and the percentage this loss represents, respectively.

<u>Figure 28.</u>
<u>Optimal portfolio Expected Return as a function of ESG Integration Index.</u>



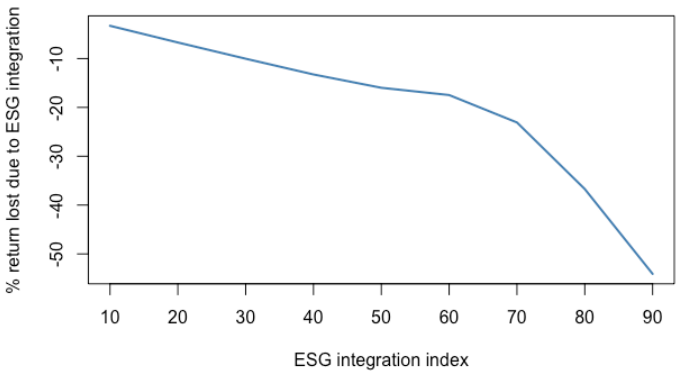
Note. Source: Refinitiv Eikon, FRED

<u>Figure 29.</u>
<u>Optimal portfolio Expected Return Absolute Loss as a function of ESG Integration Index.</u>



Note. Source: Refinitiv Eikon, FRED

<u>Figure 30.</u>
<u>Optimal portfolio Expected Return Percentage Loss as a function of ESG Integration Index</u>

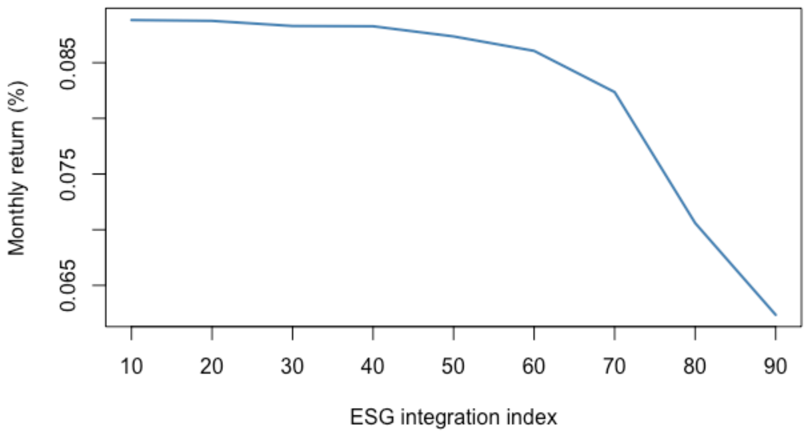


Note. Source: Refinitiv Eikon, FRED

From these graphs, we can conclude about the impact of integrating ESG considerations in terms of portfolio performance. It appears that until a certain threshold, the cost of integrating ESG preferences is low. For instance, the return cost of integrating ESG preferences at a level of 30% is around 12%, whereas the cost of integrating ESG preferences at a level of 60% is around 25%. However, after an integration level of 70%, the slope of the curve sharpens and it becomes increasingly expensive to integrate ESG preferences. For instance, the return cost of integrating ESG preferences at a level of 30% is around 90%.

Having considered the impact of ESG integration at the asset allocation stage extensively, we might now want to see how these results translate empirically, considering what has happened ex-post once the stocks returns have been realized (Figure 31). The extent to which these realized returns align with those predicted directly depends on our regression model's accuracy. The higher the predictive power, the more the ex-post curve will look similar as the one generated using predicted expected returns.

Figure 31.
Optimal Portfolio Ex-Post Realized Return as a function of ESG Integration Index.



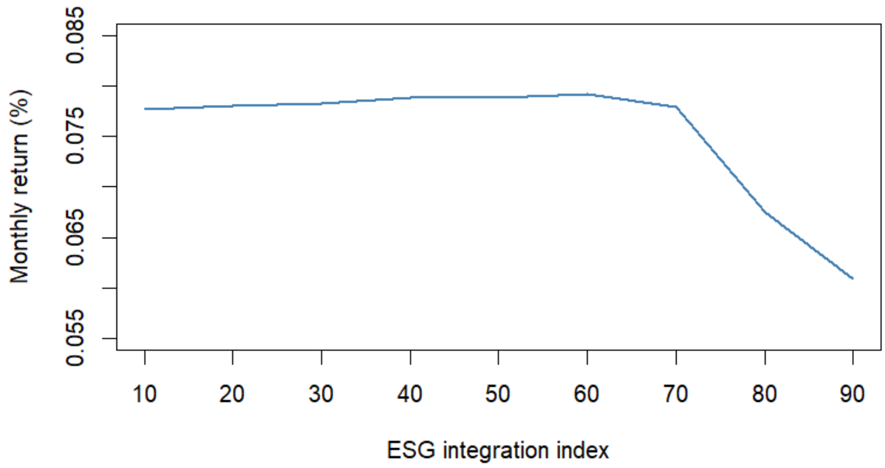
Note. Source: Refinitiv Eikon, FRED

Although the scales of returns are very different, as our predicted values may have been overly optimistic for this particular time step, we notice that the curve has a similar aspect to the curve drawn at the asset allocation time. This observation could be attributed to the fact that the MLP predictions that we are using here are the ones which implied the best investment strategy. This good performance is directly linked to the MLP model's predictive power for this particular task. The same conclusions as in the previous sub-section can thus be drawn, but with more nuance as the loss is less significant both in absolute or percentage terms for the ex-post results for extreme values in ESG integration levels.

To go even further in our study, it is interesting to measure to what extent the predictions generated by our multilayer perceptrons regression model can help a serious ESG investor to incorporate ESG preferences at a lower realized return cost as compared to more naive methods. Figure 32 illustrates the out-of-sample actual return in relation to the ESG Integration Index, which is implied by a strategy for which we employed the historical average of stock returns within the rolling window as predictions for future returns in our optimization process.

As compared to the scenario in which we use the predictions generated by the MLP model, it turns out that for a same level of ESG integration, the out-of-sample returns are systematically lower for the historical average implied strategy, thus confirming empirically the benefits of making use of machine learning models for ESG serious investors in their quest of greener portfolios.

Figure 32.
Historical Average Implied Strategy Realized Return as a function of ESG Integration Index



Note. Source: Refinitiv Eikon, FRED

5. Discussion

In the initial section of this master thesis, we presented two primary research questions that this study aims to address. Now, in this discussion section, we will revisit these questions to provide our final answers in light of the findings we explored in the results section. These conclusions will be compared with the findings from the scientific literature highlighted in the literature review. In the process, some nuances will be given to our results and we will evaluate to what extent they can be extended outside of our study. The limitations of our methodology will also be discussed extensively.

5.1. Quest for the optimal regression model for portfolio optimization

Moving from our initial conclusion that Multilayer Perceptron (MLP) implied strategy is indeed the optimal strategy among these considered, we now discuss the key factors that influenced this outcome. We'll examine the defining metrics, importance of risk profiles, and potential limitations we should be aware of while reading the results of this study.

5.1.1. Key components in addressing the research question

The first research question we wanted to address was the following: "Among a selection of regression models, determine the one which promotes the best asset allocation in a pre-defined investment universe, that yields the best out-of-sample portfolio performance based on the mean-variance optimization model".

Based on the thorough performance analysis of various regression models and their impact on asset allocation within a predefined investment universe, it is evident that the Multilayer Perceptron (MLP) strategy emerges as the optimal choice. This conclusion is driven by the robust performance of the MLP strategy across several out-of-sample measures when integrated in the mean-variance optimization model.

The MLP strategy produced the highest cumulative returns and exhibited the most impressive CAGR (12%). Exhibiting superior risk-adjusted performance for other metrics as well, such as the highest Sortino ratio of 1.91 and Information ratio of 0.15, the MLP strategy demonstrated an unrivaled capability to generate returns and provide an excellent risk-return trade-off.

While other strategies like Elastic Net Regression, Recurrent Neural Networks (RNN), Long Short-Term Memory Neural (LSTM) Networks, and Linear Regression showed respectable results in some areas, none matched the overall performance of the MLP. For instance, Elastic Net Regression presented a respectable risk-adjusted return with the highest Sharpe ratio of 2.35 but fell short with a lesser CAGR of 9%. Similarly, RNN and LSTM performed well in terms of their Treynor ratios, indicating efficient risk-adjusted returns considering their systematic risk (beta), but were unable to rival the MLP's overall performance.

Ridge and Lasso Regression strategies were found to be inferior to other methods. Ridge Regression showcased average returns and risk-adjusted metrics, while Lasso Regression lagged behind all the strategies, displaying the greatest downside risk and poorest risk-adjusted returns.

The MLP strategy thus appears as the most promising for achieving the best out-of-sample portfolio performance in the defined investment universe using the mean-variance optimization model. Nevertheless, this result is to take with caution as an investor may have a risk profile so that one or more particular metrics should have a more consequent weight in the final decision to adopt a specific regression model to predict future stock returns.

5.1.2. Limitations

First, it should be noted that the historical variance-covariance matrix was used throughout the study to allow for portfolio risk quantification. A possible enhancement of this methodology could have been the inclusion of linear and non-linear shrinkage methods, which might have potentially offered a more accurate estimate of the future variance-covariance matrix. This enhancement could have influenced the efficiency of the portfolio allocation, thereby impacting the overall performance of the regression model chosen.

Additionally, the defined investment universe was limited to the components of the Dow Jones Industrial Average (DJIA) index. This suggests that the performance of the regression models may not remain consistent with different asset selections or if applied to other indexes or broader markets. Therefore, the results of this study should be cautiously extended to other asset classes or geographical locations.

The use of the DJIA index also introduces a potential survivorship bias into the study. By including only surviving companies, the estimated performance might be artificially inflated and the effects of risk diminished, leading to a strong bias. This could influence the accuracy of the portfolio performance, and it's possible that a different regression model might prove to be optimal when considering a universe of assets that also includes companies that have failed.

A further point of consideration is the selection of macroeconomic variables. The choices in this study were made arbitrarily and could have been based on a more rigorous review of existing scientific literature. Different selections might have led to differing performances of the regression models and thus, the results may be dependent on the chosen macro-variables.

Finally, this study did not incorporate stock-specific metrics such as size (SMB) and value (HML) as in the Fama and French model (1992). Inclusion of these factors could provide a more granular insight into the risk and return characteristics of individual stocks, which in turn could have an effect on the effectiveness of the regression models. Therefore, the absence of these metrics may pose a limitation to the precision of the portfolio optimization and the generalizability of the superior performance of the Multilayer Perceptron (MLP) strategy.

5.1.3. Connections with the existing literature

While being revealing in many respects, this study contrasts with existing literature on some aspects, particularly regarding the performance of certain regression models. Particularly, the performance of the Recurrent Neural Networks (RNN) implied strategy and especially the Long Short-Term Memory Neural Networks (LSTM) implied strategy, though having shown respectable out-of-sample performance with Sharpe ratios of 2.31 and 2.01 respectively, were anticipated to perform significantly better as compared to other strategies given their architecture's inherent suitability for stock returns prediction.

The underlying principle of LSTM should theoretically make it an ideal candidate for tasks like stock return prediction. The neural network's design allows it to remember past information and utilize it in making future predictions, a characteristic that is precious for the financial markets where historical prices and returns often play a crucial role.

However, the results of our study showed a different reality. The LSTM's performance was surpassed by the Multilayer Perceptron (MLP), which denotes with the literature’s conclusions that seem to agree about LSTM's superior abilities in the context of month-ahead stock returns prediction for DJIA constituents.

One plausible explanation for this unexpected outcome might reside in the feature selection methodology employed in this study. The features were chosen on an ad hoc basis, with the optimal combination determined via trial and leading to the best 3-year cumulative return we could find. Although practical, this method might not necessarily yield the best set of features for LSTM or RNN models, potentially compromising their ability to fully exploit their capabilities for processing sequential data.

A more systematic approach, such as grid search, coupled with a time-series cross-validation technique like walk-forward validation at every prediction step, could potentially offer a more effective selection of features. However, such a method would be substantially more computing-intensive and might not be feasible given the computational resources available for the study.

Moreover, it's important to remember that conventional k-fold cross-validation could disrupt the temporal structure of the time-series data, which could potentially lead to suboptimal or misleading model performance. Thus, its application for feature selection in time-series analysis with LSTM or RNN models should be approached with caution, leading to our decision not to use these methods in this study.

5.2. Description and impact assessment of an ESG integration procedure

Having extensively explored and assessed a possible ESG integration procedure in the previous section, we move on and attempt to identify the critical elements that guided our investigation. We'll briefly come back on the visual tools and methods that helped us understand the ESG integration's implications on portfolio optimization, and the costs associated with integrating ESG preferences implied by the use of these models. We will also discuss limitations, connections with the literature review, and areas for future research.

5.2.1. Key components in addressing the research question

The second main research question that we have decided to investigate was: "For the most performant regression model, determine what happens if we introduce ESG considerations. We want to define a procedure for their integration in the investment decision-making process and understand the return cost or benefit of the inclusion of these preferences both at asset allocation stage and ex-post".

We have discussed in-depth the introduction of ESG factors into the portfolio optimization process, moving beyond traditional mean-variance optimization. To do so, we were mainly inspired by the work of Steuer and Utz (2023), who introduced a three-dimensional framework visualizing the portfolio with the maximum attainable ESG score for any combined level of risk and return. This efficient surface helps an investor visualize the trade-off between risk, return, and ESG performance.

We have made use of the concept of the ESG-efficient frontier, another visualization tool that helps an investor select a portfolio that aligns with their risk and ESG preferences. We have enhanced this model with a color scale indicating a measure of risk for any portfolio on the ESG-efficient frontier, addressing the shortcoming identified by Steuer and Utz (2023) of the initial method proposed by Perdersen et al. (2021), whom focus was only on finding the maximal the Sharpe ratio for any given of risk.

We have investigated a method to quantify ESG preferences by using the ESG Integration Index, also proposed by Steuer and Utz (2023). This index helped us understand the return level that a serious ESG investor must sacrifice to enhance the ESG performance of their portfolio for different levels of ESG integration. The results demonstrate that as the ESG Integration Index increases, the ESG score of the portfolio also increases proportionally unsurprisingly, while portfolio risk decreases for most ESG integration index values. However, there is a negative impact on portfolio performance, with portfolio returns decreasing as the ESG Integration level increases, especially after a certain threshold that lies among the highest possible values of the ESG Integration Index.

We also demonstrated empirically that using predictions generated by a machine learning model, and in particular the multilayer perceptron regression model, can lower the realized return cost associated with integrating ESG preferences into portfolio management, compared to more naive methods to predict future returns, in particular the historical average of returns among a rolling window. Overall, this study indeed provides significant insights into the potential costs associated with integrating ESG preferences into portfolio management, highlighting the potential for a trade-off between ethical considerations and portfolio performance.

5.2.2. Limitations

Undoubtedly, the major limitation in this part of our study, as previously explored, resides in our usage of static ESG scores. Despite the simplicity and convenience that this method affords, it may not accurately capture the temporal variability of ESG factors. This lack of dynamism might describe a wrong picture of these scores' impact, as they can evolve and fluctuate based on changing social, environmental, and corporate governance realities among the considered stocks.

More than ever, companies are progressively increasing their ESG efforts, implementing changes that aren't reflected in static scores. These continuous adaptations to better business practices, increased transparency and focus on sustainability goals are better captured by dynamic scoring mechanisms that reflect the true status and trajectory of ESG initiatives. Moving forward, adopting an approach that utilizes dynamic ESG scores could potentially enhance the richness of the proposed analysis and provide a more comprehensive understanding of ESG integration impacts.

Another potential limitation to our study could come from the relatively modest sample size we employed to assess the cost of integrating ESG preferences in investment decision-making. This cost was derived from the 36 periods for which we generated predictions from our multilayer Perceptron regression model. To enhance the robustness and consistency of our findings, it might have been beneficial to consider a more substantial sample size.

5.2.3. Connections with the existing literature

Drawing parallels between our empirical findings and those of Steuer and Utz (2023) study presents a challenge, due to the deviations in our respective methodologies. They conducted a somewhat analogous study on the S&P500 constituents over a five-year time horizon, yet with a few distinguishing aspects.

One crucial element of their approach was the application of a filter to exclude 50% of the S&P500 stocks that scored below the median ESG rating. This selection strategy was intended to replicate the practices of non-integrated ESG mutual funds as discussed in the literature review. In contrast, our study incorporated merely all the stocks of the DJIA index, leading to a significant distinction between the two studies. They also applied upper bounds on portfolio stock weights, a feature absent from our methodology. These divergent elements are potentially the primary source of discrepancies between our respective outcomes.

Another point worth mentioning is that Steuer and Utz (2023) study adopted the ESG score proposed by Sustainalytics. This should, by the definition of the ESG Integration Index as set by Steuer and Utz (2023) themselves, have a minimal effect on the comparability of results. Their ESG Integration

Index served as a core component of our study as well. To enhance the comparability of our findings, we selected a problem size of 100, a size also used among others by Steuer and Utz when conducting a similar analysis.

The key finding of Steuer and Utz (2023) research suggests achieving an ESG Integration Index of 10 requires a 1 basis point (bp) sacrifice, 20 requires a 5bp sacrifice, 29 necessitates a 10bp sacrifice, and 41 demands a 20bp sacrifice. Interestingly, our research revealed the ability to reach these ESG integration levels for similar basis point sacrifices. However, as Steuer and Utz (2023) did not reveal in their study what level of ESG integration is achievable for higher sacrifices in basis point, it is not possible to firmly state that we reached similar results with our slightly modified approach.

Nevertheless, we were able to corroborate the observation that incremental improvements in portfolio ESG integration become progressively costlier as we move further along the ESG Integration Index. According to the results described in both studies, a same 10% enhancement in relative ESG performance will require a higher return sacrifice when originating from the mean-variance optimal portfolio, compared to a portfolio already incorporating ESG factors to a certain level. In other terms, if Steuer and Utz (2023) had constructed a similar curve as presented in our figure 28 based on their empirical results, they would have similarly observed an increasingly steep negative slope.

6. Conclusions

In this concluding section, we reflect on the dual exploration of asset allocation strategies and the integration of ESG considerations into portfolio management. The detailed analysis of multiple regression models revealed the Multilayer Perceptrons (MLP) as the superior model, providing optimal out-of-sample performance. Furthermore, the investigation into the application of ESG factors in investment decision-making offered nuanced insights into the dynamic between ESG performance and financial returns. These findings carry significant implications for asset managers and paths for future research in sustainable investment practices. This conclusion aims to summarize the key findings, managerial and theoretical implications, and providing suggestions for future research.

6.1. Short summary

This research has examined several regression models to determine the best asset allocation strategy in portfolio management. The Multilayer Perceptrons (MLP) model has emerged as the most efficient approach, providing superior out-of-sample performance. This finding implies that MLP regression models can potentially offer the best allocation strategy in the mean-variance optimization problem, thereby enhancing portfolio performance and effectively balancing the risk-return trade-off. However, the optimal strategy may depend on specific factors, including investor risk profile, investment objectives, and market conditions.

Moreover, the research has explored the integration of ESG considerations into investment decision-making. Findings indicate that including ESG factors adds a third dimension to the conventional risk-return trade-off, enabling a more comprehensive investment decision-making framework. However, as the ESG Integration Index increases, portfolio returns decrease, indicating a trade-off between ESG performance and financial performance. Using MLP regression models can help minimize the return cost associated with ESG integration, which has significant implications for asset managers. This research encourages an ongoing review of portfolio optimization techniques and ESG scoring methodologies, as well as the need for clear communication with clients about the costs and benefits of ESG integration.

6.2. Managerial implications

Based on the above findings, there are significant managerial implications for the asset management industry, especially in the discussions to determine the optimal regression model for portfolio management, and the integration of ESG considerations into the investment decision-making process.

We have concluded that statistical predictions built upon a multilayer perceptrons (MLP) model were suggesting the best allocation strategy in the mean-variance optimization problem, yielding superior out-of-sample performance according to multiple metrics. It implies that managers should consider leveraging this strategy to optimize portfolio performance and balance the risk-return trade-off efficiently.

However, it's crucial for asset managers to understand that while MLP demonstrated the best overall results, the optimal strategy might depend on the specific risk profile and investment objectives of the investor. For instance, an investor emphasizing risk-adjusted returns might prefer the Elastic Net Regression strategy due to its highest Sharpe ratio, even though it has a lower CAGR compared to the MLP. Thus, the decision for asset managers should be context-specific and driven by the investor's risk tolerance, investment horizon, and overall investment objectives. It is also essential for managers to constantly monitor and evaluate the effectiveness of these regression models as financial markets evolve.

The second part of the research described a procedure to integrate ESG considerations into portfolio management. The research we processed suggested the integration of ESG factors into the portfolio optimization process to add a third dimension to the traditional risk-return trade-off, creating a more holistic framework for investment decision-making.

The integration of ESG preferences has shown to have an impact on portfolio performance. As the ESG Integration Index increases, there is a decrease in portfolio returns, demonstrating a trade-off between ESG performance and financial performance. This suggests that asset managers must be transparent with their clients about the potential cost of ESG integration and work with them to find the balance that best aligns with their ethical convictions and financial objectives.

Furthermore, the research indicates that utilizing a machine learning model, particularly the MLP regression model, can mitigate the return cost associated with ESG integration. This insight opens a new avenue for asset managers to leverage machine learning techniques to enhance the ESG performance of their portfolios while minimizing the adverse impact on returns.

Ultimately, the integration of ESG considerations into portfolio management is a complex process that requires a nuanced approach. Asset managers must keep up with the latest advancements in portfolio optimization techniques and ESG scoring methodologies, while also maintaining clear and open communication with their clients about the costs and benefits associated with ESG integration.

6.3. Theoretical implications

The theoretical implications of this master thesis are vast and applicable to a broad spectrum of portfolio management topics. The primary goal was set to explore various quantitative methods, particularly deep learning models, commonly used by practitioners or proposed by academic research. The overarching intention was to understand these methods' usefulness to provide superior performance against passive strategies. This aim was concretized through two main research questions:

1: "Among a selection of regression models, determine the one which promotes the best asset allocation in a pre-defined investment universe, that yields the best out-of-sample portfolio performance based on the mean-variance optimization model."

2: "For the most performant regression model, determine what happens if we introduce ESG considerations. We want to define a procedure for their integration in the investment decision-making process and understand the return cost or benefit of the inclusion of these preferences both at asset allocation stage and ex-post."

The successful investigation of these questions has led to insightful conclusions that can be used to as a key starting point for future research. The findings regarding the optimal regression model for portfolio optimization notably challenge the conclusions of the literature review. The empirical evidence demonstrated the superior performance of the Multilayer Perceptron (MLP) strategy over models that the literature finds superior like Recurrent Neural Networks (RNN), Long Short-Term Memory Neural (LSTM) Networks in the context of stock returns prediction.

Furthermore, the research provides a detailed insight into the procedure for integrating ESG considerations into investment decision-making. The thesis bridges the gap between sustainability considerations and portfolio performance, revealing a tangible cost for integrating ESG preferences into investment decision-making. The findings highlight that using predictions generated by a machine learning regression model, in particular multilayer perceptrons, can lower the realized return cost associated with integrating ESG preferences, thereby enabling to increase the propension of certain investors to integrate ESG considerations in their investment decisions.

Moreover, the identification of limitations and potential improvements in the study's methodologies indicates avenues for further research and refinements. These can be valuable for researchers who aim to advance further the research of juxtaposing the fields of portfolio optimization and sustainable investment practices.

6.4. Limitations and suggestions for future research

The results of the study pertaining to the identification of the most suitable regression model to predict stock returns highlighted the MLP implied strategy superior performance. It is the somewhat surprising underperformance of RNN and LSTM, given their theoretical advantages, that warrants further research. Future studies may need to further investigate feature selection methods and perhaps attempt a more systematic approach, despite the computational challenges, to ascertain whether the LSTM and RNN models can indeed outperform other regression models in portfolio optimization.

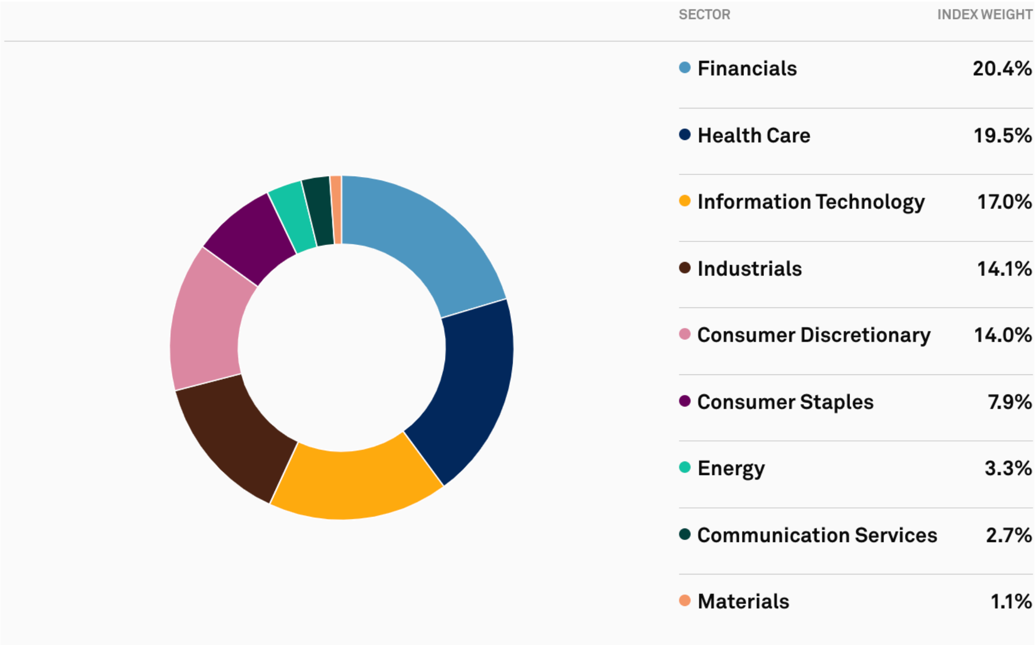
The presented research has shown a surprising result: as the ESG Integration Index increases (meaning a portfolio has higher ESG scores), the risk (as measured by expected volatility) decreases. This result defies the conventional assumption that limiting the investment universe by increasing ESG criteria would lead to less diversification and therefore more risk. Further research can be focused on validating this result with larger and diverse samples. For instance, we could extend the study to include stocks from different geographies or sectors. It would be interesting to see if this inverse relationship between ESG score and risk holds for all types of equities or if there are certain sectors, markets, or types of companies where the relationship doesn't hold or even reverses. In addition, more granular analysis could be performed by breaking down ESG scores into their Environmental, Social, and Governance components, to determine if one aspect drives these results more than the others.

7. Appendices

Table 1.
Selected Macroeconomic regressors.

FRED Ticker	Dataset Name
GDPC1	Real Gross Domestic Product
DEXUSEU	U.S./Euro Foreign Exchange Rate
DEXMXUS	U.S./Mexico Foreign Exchange Rate
DEXCAUS	U.S./Canada Foreign Exchange Rate
DEXCHUS	U.S./China Foreign Exchange Rate
DCOILWTICO	Crude Oil Prices: West Texas Intermediate (WTI)
DHHNGSP	Henry Hub Natural Gas Spot Price
DPROPANEMBTX	Propane Prices: Mont Belvieu, Texas
PNRGINDEXM	Global price of NG
APU000072610	Gasoline, All Types
PMETAINDEXM	Global price of a metal
CSUSHPINSA	S&P/Case-Shiller U.S. National Home Price Index
CPIAUCSL	Consumer Price Index for All Urban Consumers
NASDAQCOM	NASDAQ Composite Index
WILLLRGCAPPR	Wilshire US Large-Cap Price Index
WILL2500INDGR	Wilshire US Mid-Cap Growth Index
WILL2500INDVAL	Wilshire US Mid-Cap Value Index
VXDCLS	CBOE DJIA Volatility Index
M2REAL	Real M2 Money Stock
M2V	M2 Money Stock Velocity
BOPGSTB	Goods, Value of Exports, Balance of Payments Basis
FRBKCLMCIM	Industrial Capacity Utilization Rate for France
MNFCTRIRSA	Manufacturers: Inventories to Sales Ratio
BOGZ1FA895050005Q	Nonfinancial Business; Debt to Equity Ratio
MNFCTRSMSA	Manufacturing Sector: Real Output
GAFDFSA066MSFRBPHI	Philadelphia Fed ADS Business Conditions Index
NEFDFSA066MSFRBPHI	Philadelphia Fed Coincident Index
TB3MS	3-Month Treasury Bill: Secondary Market Rate
AAAFF	Moody's Seasoned Aaa Corporate Bond Yield
BAAFFM	Moody's Seasoned Baa Corporate Bond Yield
KCFSI	Kansas City Financial Stress Index
FEDFUNDS	Effective Federal Funds Rate

Table 2.
Dow Jones Industrial Average Sectors Exposure.



Note. Source: Standard & Poor's, https://www.spglobal.com/spdji/en/indices/equity/dow-jones-industrial-average/#data

Table 3.
Refinitiv ESG Score Scale Interpretation table

0 to 25	First Quartile	Scores within this range indicate poor relative ESG performance and insufficient degree of transparency in reporting material ESG data publicly.
> 25 to 50	Second Quartile	Scores within this range indicate satisfactory relative ESG performance and moderate degree of transparency in reporting material ESG data publicly.
> 50 to 75	Third Quartile	Scores within this range indicate good relative ESG performance and above average degree of transparency in reporting material ESG data publicly.
> 75 to 100	Fourth Quartile	Scores within this range indicate excellent relative ESG performance and high degree of transparency in reporting material ESG data publicly.

Note. Source: https://www.refinitiv.com/en/sustainable-finance/esg-scores#t-score-range

8. References

- Barkoulas, J.T., Baum C.F. (1996). Long-term dependence in stock returns, Economics Letters, 53(3), 253-259, https://doi.org/10.1016/S0165-1765(96)00935-4.

- Black, F., & Litterman, R. (1992). Global portfolio optimization. Financial Analysts Journal, 48(5), 28-43. https://doi.org/10.2469/faj.v48.n5.28

- Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. The Journal of Finance, 52(1), 57-82. https://doi.org/10.1111/j.1540-6261.1997.tb03808.x
- Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. Quantitative Finance, 1(2), 223-236. https://doi.org/10.1080/713665670

- David A. Hsieh (1995) Nonlinear Dynamics in Financial Markets: Evidence and Implications, Financial Analysts Journal, 51:4, 55-62.

- D'Hondt, C., De Winne, R., Ghysels, E. & Raymond, S. (2020). Artificial intelligence alter egos: Who might benefit from robo-investing? Journal of Empirical Finance, 59, pp. 278 299. https://doi.org/10.1016/j.jempfin.2020.10.00

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

- European Union. (2014). Directive 2014/65/EU of the European Parliament and of the Council of 15 May 2014 on markets in financial instruments and amending Directive 2002/92/EC and Directive 2011/61/EU. Official Journal of the European Union, L 173(1), 349-496.

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

- Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. European Journal of Operational Research, 270(2), 654-669. https://doi.org/10.1016/j.ejor.2017.11.054

- Friedman, J., Hastie, T., & Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. Biostatistics, 9(3), 432-441. DOI: 10.1093/biostatistics/kxm045

- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

- Hodnett, K., & Hsieh, H. (2012). Capital Market Theories: Market Efficiency Versus Investor Prospects. International Business & Economics Research Journal, 11(8), 849-862. https://doi.org/10.19030/iber.v11i8.7163

- Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. Technometrics, 12(1), 55-67. https://doi.org/10.2307/1267351
- LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. Nature 521, 436–444 (2015). https://doi.org/10.1038/nature14539

- Ledoit, O., & Wolf, M. (2004). Honey, I shrunk the sample covariance matrix. The Journal of Portfolio Management, 30(4), 110-119. https://doi.org/10.3905/jpm.2004.110

- Lipton, Z. C. (2015). A critical review of recurrent neural networks for sequence learning. Retrieved from: https://www.researchgate.net/publication/277603865_A_Critical_Review_of_Recurrent_Neural_Networks_for_Sequence_Learning

- Maciel, L. S., & Ballini, R. (2009). Design a neural network for time series financial forecasting: Accuracy and robustness analysis. Instituto de Economia (IE), Universidade Estadual de Campinas (UNICAMP). https://www.cse.unr.edu/~harryt/CS773C/Project/895-1697-1-PB.pdf

- Markowitz, H. (1952). Portfolio selection. The Journal of Finance, 7(1), 77–91. https://doi.org/10.1111/j.1540-6261.1952.tb01525.x

- Marszk, A., & Lechman, E. (2023). Sustainable investing: Socio-economic impacts of exchange-traded funds. Elsevier.

- McIntyre, C., Bartletta, S., Bhattacharya, I., Carrubba, J., Frankle, D., Frazao, M., Heredia, L., Mörstedt, H., Newsom Reeves, K., Palmisani, E., Pardasani, N., Reeves, M., Schulte, T., Sheridan, B., Slack, B., & Zhu, K. (2022, May). BCG Global Asset Management 2022—20th Edition: From tailwinds to turbulence. Retrieved from https://www.bcg.com/publications/2022/tailwinds-to-turbulence-for-global-assets-under-management

- McLean, R.D., & Pontiff, J. (2016), Does Academic Research Destroy Stock Return Predictability? Journal of Finance, 71, 5-31. http://dx.doi.org/10.2139/ssrn.2080900

- Nielsen, M. A. (2015). Neural Networks and Deep Learning. Determination Press.

- Oikonomou, I., Platanakis, E., & Sutcliffe, C. (2017). Socially responsible investment portfolios: Does the optimization process matter? British Accounting Review, 50(4), 379–401. https://doi.org/10.1016/j.bar.2017.10.003

- Olah, C. (2015, August 27). Understanding LSTM Networks. Retrieved from http://colah.github.io/posts/2015-08-Understanding-LSTMs/

- Pedersen, L.H., Fitzgibbons, S., & Pomorski L. (2021). Responsible investing: The ESG-efficient frontier. Journal of Financial Economics, 142(2), 572-597. https://doi.org.10.1016/j.jfineco.2020.11.001

- Reuters Staff. (2023, March 8). DowDuPont to separate material science unit on April 1. Reuters. Retrieved May 14, 2023, from https://www.reuters.com/article/us-dowdupont-divestiture-idUSKCN1QP1BW

- Refenes, A. N., Burgess, A. N., & Bentz, Y. (1997). Neural networks in financial engineering: a study in methodology. IEEE Transactions on Neural Networks, 8(6), 1222-1267.

- Rumelhart, D., Hinton, G. & Williams, R. (1986). Learning representations by back-propagating errors. Nature, 323, 533–536. https://doi.org/10.1038/323533a0

- Steuer, R.E. & Utz, S. (2023). Non-contour efficient fronts for identifying most preferred portfolios in sustainability investing, European Journal of Operational Research, 306 (2), 742-753. https://doi.org/10.1016/j.ejor.2022.08.007

- Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. Journal of the Royal Statistical Society: Series B (Methodological), 58(1), 267-288. https://doi.org/10.1111/j.25176161.1996.tb02080.x

- Tibshirani, R. (2011). Regression shrinkage and selection via the Lasso: A retrospective. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 73(3), 273-282. https://doi.org/10.1111/j.1467-9868.2011.00771.x

- Utz, S., Wimmer, M., Steuer, R. E., & Hirschberger, M. (2014). Tri-criterion inverse port- folio optimization with application to socially responsible mutual funds. European Journal of Operational Research, 234(2), 155–164. DOI: 10.1016/j.ejor.2013.07.024

- Zou, H., & Hastie, T. (2005). Regularization and variable selection via the Elastic Net. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 67(2), 301-320. https://doi.org/10.1111/j.1467-9868.2005.00503.x

Executive summary

This master thesis, through the integration of deep learning models and the inclusion of ESG preferences, presents a comprehensive approach to portfolio management that can serve as a reference for sustainable investing in the future. By providing insights into the complexities and trade-offs involved in portfolio optimization and ESG integration, this research facilitates more informed decision-making for portfolio managers, enabling them to balance performance goals with ethical considerations effectively. The integration of these key components reinforces the thesis's primary goal and substantiates its significant contribution to the fields of portfolio management and sustainable investing.

Building upon these findings, the research discovered the Multilayer Perceptrons (MLP) model to be an especially powerful tool in asset allocation. Its robust performance promises to increase the efficiency of portfolio management, effectively managing the risk-return trade-off. However, it is of the utmost importance for asset managers to align the choice of regression models with specific investor profiles and market conditions, highlighting the need for flexibility in strategy implementation.

In relation to ESG integration, our study underscored the importance of balancing ethical considerations with financial performance. The trade-off observed between ESG performance and portfolio returns prompts asset managers to approach ESG integration with caution. However, the research also demonstrates that the application of strategies implied by machine learning models like the MLP can potentially mitigate the return cost associated with ESG integration, paving the way for more sustainable investing while limiting the compromise on returns.

This research not only offers practical insights for portfolio managers but also opens up exciting paths for future exploration. With the ongoing evolution of financial markets and sustainable investing, this study underscores the importance of continual review and innovation in portfolio optimization techniques and ESG integration. As such, this master thesis contributes to the current discourse in the field of portfolio management and sustainable investing, serving as a key starting point for more advanced and comprehensive approaches to sustainable portfolio management in the future.

[Word count = 22,459]