UCLouvain

LOUVAIN
School of Management

Louvain School of Management

Design of ESG-compliant portfolios: a minimum divergence approach

Author(s): HALLEUX Laura
Supervisor(s): VRINS Frédéric

Academic year 2021-2022
Dissertation for the master of Business Engineer
Master [120]: Business Engineering – Financial Engineering
Daytime schedule

Abstract

By taking into account the key subject of our society, the climate change, and by using the minimum divergence approach introduced by Lassance & Vrins (2022), in this master thesis, we will design ESG-compliant portfolios following two different strategies: (i) implement basic portfolios such as equally-weighted portfolios, Sharpe ratio portfolios and minimum-variance portfolios on a restricted universe considering ESG criteria; and (ii) implement minimum-divergence approach on portfolios with a good ESG score by targeting a Gaussian distribution with the two first moments of benchmark portfolios on an unrestricted universe. Thanks to the in-sample and out-of-sample analysis, we will compare three types of portfolios to each other: the benchmark portfolios, the ESG traditional ones and the minimum divergence portfolios. The goal of this master thesis is to see if it is possible while targeting the distribution of a certain portfolio on an unrestricted universe to improve the performance measures of an ESG-compliant portfolio. We concluded that in-sample it was generally possible for the minimum divergence portfolios to outperform the ESG traditional portfolios except for the Sharpe ratio strategy in terms of skewness, kurtosis and adjusted Sharpe ratio. For the out-of-sample performance, the results were poor for dataset with more assets because of the estimation errors and the results for the dataset with less assets were significant only with a small estimation window.

"All growth starts at the end of your comfort zone"

- Tony Robbins

Acknowledgements

Before anything else, I would really and sincerely like to thank my supervisor, Prof. Frédéric Vrins, who accompanied me throughout this thesis. Since the beginning of my thesis, starting with the acceptance to be my supervisor, passing by the proposal of topics, the kindness during my learning of the main method used in my thesis, the advice, the time and the patience he showed for the implementation of my code and finally ending by the careful and attentive reviewing of my thesis, he has always been present for each important step of this work.

I would also like to thank Prof. Nathan Lassance who was always very reactive when I had technical questions about the method to be implemented or about risk measures. He was always willing to set up a call if it was necessary for me. His availability and time were precious.

People I will always thank are my family members. Let's start with my parents who since the beginning of my studies have always been there for me and without whom I would never have been able to consider studying and therefore submit this thesis. Their support has always been unfailing, their encouragement always at the right moment and their arms always open and comforting. Thank you to them for being there in the good and bad times, for believing in me from the beginning to the end and for their patience and kindness.

I would also like to thank my little brother who always offered me his help to think about the logic of my code, to have always been there to support me in his own way and to have undergone my mood changes. Thanks to him for having done everything to take my mind off my studies, the revision periods and the writing of this work.

Another person I would like to thank is my boyfriend who has understood and accompanied me throughout these studies and my thesis. An impressive patience and a bearing love that gave me the strength and courage to continue this important journey of my life.

I would also like to say thank you to my close friends who made these years of study an unforgettable memory. They were the first to go for a drink to discuss our different dissertations, to exchange advice or simply to change our minds.

And finally, thank you to you, dear reader, who take the time to read and consider my work. I hope you will enjoy it and that you will find it interesting.

Table of contents

List of Figures

List of Tables

Introduction

The climate change is currently a key subject in our society. Following the objective of Paris Agreement to keep the global temperature increase below 2°C, Europe wants to become the first climate-neutral continent of the world by 2050. Therefore, the European Commission establishes different regulations that will be more and more restrictive for investors. Concerning the regulation on banking and finance, the objective of the European Commission is to transform the current finance into sustainable finance which is a process of considering the environmental, social and governance factors in the investment decision-making process. The stakeholders demand changes then and they give more importance to sustainability in their decision-making process. Several authors indicate that ESG investments can outperform on long-term basis investment that do not consider ESG criteria. Therefore, it begins to be a necessity to take into account the ESG criteria and investors want to implement that into their portfolio strategy. It exists different ways to implement these criteria such as assets exclusion or penalization. Thanks to this implementation, portfolios computed could be ESG-compliant.

The portfolio-selection is the basis way to allocate wealth into several assets. As Lassance & Vrins (2022) noted, there are three important steps in order to do a correct wealth allocation. The first one is to find an objective function that will translate the preferences of investors. The second one is to find optimal weights based on this objective function that is maximized or minimized depending on the preferences. And the last one is performing an out-of-sample analysis.

In order to find the optimal weights, Markowitz (1952) proposed to solve a quadratic problem which is to maximize the returns of a portfolio while penalizing for risk. The penalization is represented by the risk aversion of the investors. If the investor does not like the risk, the penalization will be high in order to reduce the risk as much as possible and if the investor likes the risk, the penalization will be low. This quadratic problem will give rise to the efficient frontier then the mean and variance of all efficient portfolios. Based on this efficient frontier, the portfolio maximizing the Sharpe ratio is the portfolio that investors want to achieve. In order to reach this portfolio, it will be estimated by the sample mean-variance portfolio where the mean is estimated by the sample mean and the covariance matrix by the sample covariance matrix.

However, the estimation of the mean returns is difficult due to the fact that the mean-variance portfolio is very sensitive to the sample mean (Chopra & Ziemba, 1993) and also the variance of the sample mean is very high (Black, 1993). Therefore, the mean-variance portfolio is very sensitive to estimation error in the inputs which deteriorates the out-of-sample performance.

In order to improve the estimation of mean-variance portfolio, assuming that the sample mean of all assets is the same can be a solution and it corresponds to the minimum-variance portfolio. However, the minimum-variance portfolio is again sensitive to estimation errors which is due to the covariance matrix estimation. Therefore, many authors suggest more robust estimator for the covariance matrix in order to improve the out-of-sample analysis such as Ledoit & Wolf (2017) or DeMiguel & Nogales (2009). Or Meucci (2009) and Lassance et al. (2021) who introduced a different way of estimating the covariance matrix through dimension reduction.

However, Lassance & Vrins (2022) introduced a novel framework which does not focus on the improvement of the estimation for the sample mean or the sample covariance matrix to find optimal weights but this framework will focus on the first step, the translation of investors' preferences in the objective function. Indeed, they highlight the fact that the returns do not follow a Gaussian distribution and that investors are also interested in higher moments such as skewness and kurtosis. Then, they concluded that "the mean-variance preferences is unrealistic" (Lassance & Vrins, 2022). Therefore they propose to translate investors' preferences by a target-return distribution instead by an utility function. And the optimization problem becomes the minimization of the distance between the portfolio-return distribution and the target-return distribution of investors.

Based on the actuality of our society and the new framework proposed by Lassance

& Vrins (2022), this master thesis has as objective to see if it is possible to achieve better results for ESG-compliant portfolios concerning its higher moments.

After performing an in-sample and out-of-sample analysis using a 10-years estimation window and an exclusion rate of 20 for the ESG risk score, we will see if the sample size of estimation window for rolling window method has an impact on out-of-sample performance and if the conclusions remain the same when the exclusion threshold is changed.

In **Chapter 1**, the fundamental Modern Portfolio Theory will be quickly recalled to easily understand the foundation of this master thesis.

In **Chapter 2**, the traditional portfolios such as minimum-variance, Sharpe ratio and equally-weighted portfolios will be presented because they are going to be used for the construction of different portfolios later. The two first moments of each will be used as the two first moments of a Gaussian distribution that will be helpful as target distribution for the minimum divergence strategy (see chapter 4).

In **Chapter 3**, the reasons of why considering ESG criteria will be exposed and also how it will be possible to integrate these ESG criteria in the different construction of portfolios in order to build ESG-compliant portfolios.

In **Chapter 4**, the concept of the minimum divergence strategy is explained and practical aspects related to its implementation are discussed.

In **Chapter 5**, the data and methodology will be precisely explained. The results and the further analysis in-sample and out-of-sample will be performed in this chapter. And we will evaluate the impact of the estimation window as well as the effect of the change for the exclusion threshold.

And finally, in **Chapter 6**, we will conclude this master thesis by reminding the main results. We will also explain the limitations of this work and the methods used. And we will finish by explaining which are the tracks which would seem interesting to us to explore in order to go further.

Chapter 1

Modern Portfolio Theory

In 1952, Markowitz (1952) introduced a revolutionary framework known as the Modern Portfolio Theory (MPT) aiming to solve the asset allocation problem: maximizing returns with a certain level of risk. The learning outcome of this theory is diversification. It is essential to consider the impact that each asset has on portfolio's returns and risks instead of considering individually the asset's returns and risks in order to maximize the overall return or minimize the overall risk of the portfolio. Then, it is a theory that has for objective to find the portfolio that maximizes an agent's utility, formulated as a trade-off between mean and variance of returns, whose weighting of these depends on his/her risk aversion.

Suppose an investor with a certain wealth wants to invest in several assets in an universe of N assets. Therefore, he/she needs to decide the proportion of his/her wealth he/she is going to assign to each asset. This proportion allocated to the *i-th* asset in the proportion is called the weight i and noted w_i. The portfolio is represented by a N-dimensional vector of weights $w = (w_1, w_2, ..., w_N)$ and in order to invest the entire investor's wealth (no more, no less), the portfolio weights sum to one and belong to the set of reals (Lassance, 2019).

$$W := \{w \in \mathbb{R}^N | 1'w = 1\} \tag{1.1}$$

The asset returns X are also a N-dimensional vector $X = (X_1, X_2, ..., X_N)$. The asset mean-return vector $\mu = (\mu_1, \mu_2, ..., \mu_N)$ and covariance matrix Σ are defined as follows:

$$\mu_i := \mathbb{E}[X_i] \tag{1.2}$$

$$\Sigma_{ij} := \mathbb{E}[(X_i - \mu_i)(X_j - \mu_j)] \tag{1.3}$$

We always assume that Σ is of full rank and thus invertible.

The portfolio expected rate of return is the weighted sum of the assets' expected rates of return in the portfolio.

$$X_P = \sum_{i=1}^{N} w_i X_i := w'X \tag{1.4}$$

with mean and variance

$$\mu_P = \sum_{i=1}^{N} w_i \mu_i := w'\mu \tag{1.5}$$

$$\sigma_P^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} w_i w_j \sigma_{i,j} := w'\Sigma w \tag{1.6}$$

The investor has the possibility to choose among several combinations of mean μ_P and variance σ_P^2 based on his choice for the portfolio w. However, the investor would prefer to select an efficient combination, the one with minimum variance σ_P^2 for a given mean μ_P or with maximum mean μ_P for a given variance σ_P^2 (Markowitz, 1952).

$$\arg\max_{w \in W} \mu_P \quad s.t. \quad \sigma_P \leq \sigma_0 \quad or \quad \arg\min_{w \in W} \sigma_P \quad s.t. \quad \mu_P \geq \mu_0 \tag{1.7}$$

These combinations form the efficient frontier in figure 1.1. The efficient frontier is the line in the (σ,μ) plan associated with the set of portfolios solving these problems, playing with σ_0 or μ_0, respectively. And when we introduce a risk-free asset r_f, the efficient frontier becomes a line known as the capital market line.

Another possibility to obtain efficient portfolios is by choosing a value for the risk aversion instead of fixing σ_0 and μ_0. In order to formulate this investment problem, the optimization is defined via the maximization of the quadratic utility function:

$$\arg\max_{w \in W} U(w) = \arg\max_{w \in W} \mu_P - \frac{\lambda}{2}\sigma_P^2 \tag{1.8}$$

where λ is the relative risk aversion parameter and is set exogenously by the investors.

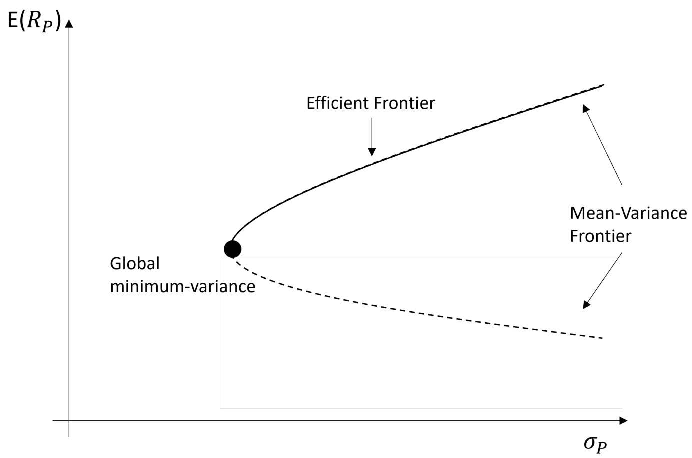


Figure 1.1: Efficient Frontier illustration

Chapter 2

Benchmark Portfolios

2.1 Minimum Variance

The global minimum variance is the portfolio with the smallest variance on the efficient frontier. And this portfolio is defined as follows:

$$w_{MV} := \underset{w \in W}{\arg\min}\ \sigma_P^2 \tag{2.1}$$

$$w_{MV} = \frac{\Sigma^{-1}1}{1'\Sigma^{-1}1} \tag{2.2}$$

2.2 Sharpe Ratio

The maximum Sharpe Ratio portfolio or the mean-variance portfolio is the portfolio maximizing the Sharpe ratio.

The portfolio is defined as follows:

$$w_{MSR} = \underset{w \in W}{\arg\max} \frac{w'\mu}{\sqrt{w'\Sigma w}} \tag{2.3}$$

And if there is only the following constraint: $W := \{w \in \mathbb{R}^N | 1'w = 1\}$, the solution is:

$$w_{MSR} = \frac{\Sigma^{-1}\mu}{1'\Sigma^{-1}\mu} \tag{2.4}$$

2.3 Equally weighted

The equally weighted portfolio is a portfolio that assigns the same importance to each asset and which is defined as follows:

$$w_{EW} = (\frac{1}{N}, ..., \frac{1}{N}) \tag{2.5}$$

2.4 Risk Parity

Diversification is essential in portfolio selection. The mean-variance portfolio is optimally diversified in the sense that it has the lowest risk (variance) for a given mean return. However, it does not work well out-of-sample and it has extreme weights. The equally weighted portfolio is optimally diversified in terms of weights because it gives the same weight to all assets. Then, it is well diversified in terms of capital because it puts the same percentage of our capital to all the assets. However, it is a naive definition of diversification as the assets are not all the same in general. In order to diversify our portfolio in terms of risk there is the risk parity portfolio (Qian, 2022) for which all the assets contribute equally to the risk of the portfolio. If this diversification is done on assets, it is called the asset-risk parity (Maillard et al., 2010). Moreover, if it is done on uncorrelated factors underlying asset returns it is called factor-risk-parity portfolio (Lassance, 2019).

2.4.1 Asset-risk parity

Markowitz theory provides portfolios that maximize the quadratic utility for some level of risk aversion coefficient, but this is not the only way to proceed. The 1/N, for instance, is just a naive way to obtain diversified portfolios. However, it does not take into account the specificity of each asset. Another approach would consist in allocating the wealth in such a way that each asset contributes equally to the overall risk of the portfolio. Considering volatility as risk-measure, asset risk parity consists in balancing the portfolio such that each asset contributes equally to the total portfolio variance. By doing so, highly-volatile assets will receive less weights than stable ones.

2.4.2 Factor-risk parity

With this asset-risk parity method, there is a problem when the asset returns are correlated because the diversifying risk across assets is suboptimal. For example, if in an investment there is 5 assets where 4 are exposed to an certain factor and the last one is exposed to another factor, then the asset-risk-parity portfolio will be more exposed to the first factor and therefore is not well diversified. Then, it is needed to change risk-parity approach where the portfolio-return risk is equally spread among uncorrelated factors rather than correlated assets.

Also, there is an important problem in portfolio selection: the "curse of dimensionality" (Lassance & Vrins, 2021). The larger the number of assets and the higher the order of the comoment, the more parameters to estimate. Therefore, the dimension reduction could help for this problem.

2.4.2.1 PCVP Portfolio

Meucci (2009) takes the principal components (PCs) as uncorrelated factors and it is called the PC-variance-parity portfolio. Then, each of the PCs contributes equally to the portfolio-return volatility. This method will help to reduce dimension because it will describe the dependence structure between the N asset returns by K $\leq$ N factors (Lassance & Vrins, 2021).

This method is also used in order to have a more robust estimator for the covariance matrix because due to the dimension reduction, there are less parameters to estimate and therefore, less chance to be exposed to estimation errors.

2.4.2.2 ICVP Portfolio

The Principal Components (PCs) are very useful for dimension reduction. However, decorrelation is not as strong as independence (except in the multivariate Gaussian case, where the concepts of correlation and dependence actually coincide) and also, the number of comoments to estimate will remain high. That is why Lassance et al. (2021) proposed the IC-variance-parity portfolio (ICVP portfolio) which will try to find factors that are as independent as possible and also will help to reduce the number of comoments needs to be estimated.

2.5 Consider higher moments

As mentionned earlier, in the mean-variance framework, an investor wants to maximize his expected utility function:

$$w^* = \arg\max_{w \in W} \mathbb{E}[G(X_P)] \tag{2.6}$$

where $X_P = \sum_{i=1}^{N} w_i X_i$. And the expected utility function can be approximated by a Taylor's expansion:

$$\mathbb{E}[G(X_P)] = G(\mathbb{E}[X_P]) + \frac{G''\mathbb{E}[X_P]}{2}(\mathbb{E}[X_P - \mathbb{E}[X_P]])^2 + \sum_{k=3}^{\infty}(\mathbb{E}[X_P]\frac{G^K(\mathbb{E}[X_P])}{k!}(\mathbb{E}[X_P - \mathbb{E}[X_P]])^k \tag{2.7}$$

where $(\mathbb{E}[X_P - \mathbb{E}[X_P]])^k$ is the kth-order centered moment. And because the utility function of investors is considered as a quadratic utility function, one can ignore the higher-moments.

It is important to note that MPT can be considered as valid if either (i) the returns are multivariate Gaussian, because in this case any combination of those assets will end up in a portfolio whose return is normally distributed, or if (ii) investors only care about the first two moments.

However, as Cont (2001), Xiong & Idzorek (2011) observed, the asset returns distribution is not a Gaussian one. Indeed, the Gaussian distribution has a third moment (skewness) equal to 0 and a fourth moment (kurtosis) equal to 3 but Cont (2001) exposes different stylized facts such as extreme events tend to cluster in time. This fact is supported by Xiong & Idzorek (2011), they argue that the distribution of asset returns is more leptokurtic, or fatter tailed, than normal distribution. The kurtosis is then higher than its value in a Gaussian distribution. Another stylized fact is that the asset returns distribution is not symmetric (Cont, 2001). This fact is also supported by Xiong & Idzorek (2011). They note that asset returns distribution is asymmetrical, the distribution is skewed to the left of the mean expected value. The skewness is then negative and not equal to zero as suggested in the Gaussian distribution.

Moreover, Xiong & Idzorek (2011) insist on the fact that skewness and kurtosis are important for investors. They are concerned about significant losses, then the downside risk.

2.6 Risk Measurement for portfolios

It is important for investors to measure risks they are potentially facing.

2.6.1 Value-at-Risk

The loss that can occur over a given term, at a particular confidence level, due to market risk is referred to as Value-at-Risk (VaR) (Bazak & Shapiro, 2001). It is a measure of the risk of loss for investments. It estimates how much we can lose with a given probability. Mathematically, the VaR is the (1-α)-quantile of the distribution. For instance, if the VaR for $\alpha = 1\%$ is 100 000, there is a one percent probability that the loss exceeds 100 000.

VaR as a function of α is defined as:

$$VaR(\alpha) = -F^{-1}(\alpha) \tag{2.8}$$

where $F^{-1}(\cdot)$ is the quantile function associated to the cumulative distribution function $F(\cdot)$. If $F(\cdot)$ is replaced by the Gaussian distribution function, the Gaussian VaR is obtained:

$$GVaR(\alpha) = -\mu_P - \sigma_P z_\alpha \tag{2.9}$$

where z_α is the gaussian quantile at confidence level α (Boudt et al., 2008).

2.6.2 Modified Value-at-Risk

However, Gaussian VaR takes only into account the first two portfolio moments and as mentionned in section 2.5, the asset returns are skewed and fat tailed. Boudt et al. (2008) add that incorporating the asymmetry and heavy tails "should lead to more accurate risk forecasts". Then, Zangari (1996) proposes another estimator that corrects GVaR for skewness and excess kurtosis and calls it the Modified VaR (MVaR).

The Modified Value-at-Risk is defined as follows:

$$MVaR(\alpha) = GVaR(\alpha) + \sigma_P[-\frac{1}{6}(z_\alpha^2 - 1)\zeta_P - \frac{1}{24}(z_\alpha^3 - 3z_\alpha)\kappa_P + \frac{1}{36}(2z_\alpha^3 - 5z_\alpha)\zeta_P^2)] \tag{2.10}$$

This expression can also take the following form:

$$MVaR(\alpha) = -\mu_P - \sigma_P(z_\alpha + \frac{1}{6}(z_\alpha^2 - 1)\zeta_P + \frac{1}{24}(z_\alpha^3 - 3z_\alpha)\kappa_P - \frac{1}{36}(2z_\alpha^3 - 5z_\alpha)\zeta_P^2) \tag{2.11}$$

where z_α is the Gaussian quantile at confidence level α, ζ_P is the portfolio skewness and κ_P is the portfolio excess kurtosis.

2.6.3 Sharpe Ratio

Sharpe (1994) introduces a ratio called Sharpe Ratio (SR) that helps investors to understand the return of an investment compared to its risk. Its expression is the following

$$SR = \frac{X_P - r_f}{\sigma_P} \tag{2.12}$$

where r_f is the risk-free rate, X_P the portfolio return and σ_P the standard deviation of the portfolio excess return. If the Sharpe ratio is below 0, it is not good because the portfolio's performance is inferior to that of a risk-free investment. If the SR is between 0 and 1, it is also not good because the excess return is insufficient to compensate for the investor's risk. And if the SR is above 1, it means that the excess return is positive and more than sufficient to compensate for risk. One important thing to know is that the Sharpe Ratio is a risk-adjusted performance measure that focuses on the first two moments, and disregards the others.

2.6.4 Adjusted Sharpe Ratio

However, as mentionned in section 2.5, the asset returns are skewed and fat tailed. Then, the Sharpe Ratio needs to be adjusted to take into account higher moments. Then, Pézier (2004) suggests the Adjusted Sharpe Ratio that not only does it dislike variance, but it also dislikes negative skewness ($\zeta_P < 0$) and positive excess kurtosis ($\kappa_P - 3 > 0$).

The expression of the Adjusted Sharpe ratio (ASR) is defined by

$$ASR = SR[1 + (\frac{\zeta_P}{6})SR - (\frac{\kappa_P - 3}{24})SR^2] \tag{2.13}$$

where SR is the Sharpe Ratio, ζ_P is the portfolio skewness and κ_P is the portfolio kurtosis.

Chapter 3

ESG Criteria

Before going into details of the ESG criteria, it is important to have the big picture of why considering ESG criteria is important.

3.1 ESG criteria

ESG is the abbreviation for Environmental, Social and Governance. Environmental factors are factors linked to the natural world (carbon emissions, deforestation, energy, water, ...). Social factors are the ones affecting life of human beings (labor standards, diversity, human rights, ...). And finally, Governance factors are factors linked to the functioning of the business model (board composition, bribery and corruption, audit committee structure, ...) (European Commission, 2022e).

These factors are more and more used by investors to identify material risks and growth opportunities.

3.2 Definition of impact

More and more investors want to make the difference instead of only making money. We can qualify them as impact investors (Brest & Born, 2013). These investors look for ways to make financial investments that will have a positive social or environmental impact.

Brest & Born (2013) define the practice of impact investing as "actively placing capital in enterprises that generate social or environmental goods, services, or ancillary benefits such as creating good jobs, with expected financial returns ranging from the highly concessionary to above market".

3.3 Why considering impact?

There are 4 main reasons to explain why it is important to consider impact.

3.3.1 Global Challenges

In December 2015, at the Paris climate conference, the Paris Agreement has been adopted. This agreement is the first to impose legal rules on the whole world to limit global warming. Its objective is not to exceed a warming of more than 2°C and then to do everything to limit this warming to 1.5°C (European Commission, 2022a).

The United Nations then set up 17 sustainable development goals (SDGs) to meet the requirements of this agreement. And so, the European Union established in December 2019 the EU Green Deal which will meet the objectives of the Paris Agreement and the SDGs. Its primary objective is to significantly reduce emissions across the European Union by 2030 and reach net zero emissions by 2050 (European Commission, 2022c). The European Union has put in place a specific action plan on Financing Sustainable Growth that covers several elements such as the EU Taxonomy, Sustainable Finance Disclosure Regulation (SFDR) and Corporate Sustainability Reporting Directive (CSRD).

3.3.2 EU Regulations

As explained above, European Union has established several regulations in order to achieve its objective and among these regulations, three main ones are EU Taxonomy, SFDR and CSRD. These regulations will impose more rules for companies and financial market participants related to their reporting and more specifically on their non-financial reporting. They will need to be more transparent about their activities and about their ESG considerations.

The EU Taxonomy is a regulation that applies to companies and financial market participants. It is a classification system that will decide if an activity is sustainable or not. The EU defines six environmental objectives: climate change mitigation (1), climate change adaptation (2), sustainable use and protection of water and marine resources (3), transition to a circular economy (4), pollution prevention and control (5) and protection and restoration of biodiversity and ecosystems (6). Based on technical screening criteria clearly defined for high climate impact sectors (sectors that are most likely to have a large impact on climate change), companies can judge if their economic activities contribute to one of the six environmental objectives while do not significantly harm the five other environmental objectives. If the companies judge that their activities contribute to one environmental objective and do not harm the other environmental objectives, the economic activities need to meet the minimal social safeguards (in the OECD guidelines, UN Guiding Principles on Business and Human Rights, and Labour Rights conventions). Finally, companies can compute, based on their economic activities that meet all the conditions above or not, the percentage of its alignment to the EU taxonomy. Financial market participants also have the obligation to report on its portfolio's taxonomy-alignment based on the taxonomy-alignment percentage of the companies in its portfolio and the weights assigned to these different companies (European Commission, 2022d).

The SFDR is applied only on financial market participants (FMPs) and financial advisers. They need to classify their financial products between two different articles: Article 8 or Article 9. Article 8 corresponds to financial products that promote environmental and/or social characteristics. Article 9 corresponds to financial products that have sustainable investment as its investment objective. Based on this classification, FMPs need to report the strategies put in place to become more sustainable, their targets and the evolution of their social and environmental KPIs (European Commission, 2022f).

And finally, the CSRD is applied only on companies. They need to do a non-financial reporting in which they are required to publicly report on their social and environmental KPIs (European Commission, 2022b).

3.3.3 Returns

Several studies prove that based on a long-term period, investing in the ESG outperforms investments that do not include these criteria.

Indeed, Verheyden et al. (2016) conclude that screening ESG provides a higher annual performance and even indicate that the volatility and CVaR are lower than the unscreened universe.

Dorfleitner et al. (2018) demontrate that asset manager can generate abnormal returns if they invest in companies with high CSR and if they hold stocks for a longer period. At the MSCI ESG Ratings Webinar organized by ASX (2021), Morgan Ellis, from MSCI, insists on the fact that the use of ESG risk data can help to reduce risk and increase returns.

And finally Broadstock et al. (2021) confirm that ESG performance mitigates financial risk during financial crisis, such as COVID-19 and that high-ESG portfolios outperform low-ESG portfolios.

However, other studies and authors do not agree with these points of view. Some argue that the abnormal returns obtained by investing in ESG are not due to the ESG nature of the investment but due to factors that have long been known to boost returns (Johnson, 2021). The situation is not clear and many articles still contradict each other on this aspect of returns of a portfolio taking into account ESG criteria.

3.3.4 Stakeholders demand

There are a lot of stakeholders that start to give a great importance to sustainability in their decision-making process. In the following paragraphs, some examples will be presented.

Customers would like to change their consumption habits to reduce their impact on climate change (Dillon, 2020).

Employee satisfaction is higher when companies have good ESG scores. The fact that considering ESG criteria and incorporating that into the practices of company can attract new workers (Bailey et al., 2020).

Bankers provide a lower cost of credit for companies who take into account ESG criteria and in which practices and programs prove that ESG criteria are in their values (Clancy, 2021).

The most of financial advisors already offer or plan to offer quickly solution incorporating ESG investments (Bates, 2021).

3.4 ESG criteria Implementation

In the literature there are many ways to implement ESG criteria in the portfolio optimization. Some decide to implement the ESG criteria separately, i.e., E, S, and G, and others implement the criteria holistically. Hong & Kacperczyk (2009), for example, uses a metric known as "sin stocks," which are publicly traded enterprises that deal in alcohol, cigarettes, and gambling, as a proxy for ESG S. For the overall ESG score, the scores provided by ESG Rating agencies are very often used as a proxy (Pedersen et al., 2021).

Once the way of representing the ESG criteria is chosen, it has to be implemented in the model. There are many different ways to implement these criteria such as imposing constraints and penalties.

When we impose a constraint, it means that we will exclude assets based on their ESG score. A limit on the ESG score will be imposed and all assets above it will be excluded from the calculation of the optimal weights.

When a penalty is applied, it is added to the objective function and penalizes heavily or not, depending on the weight assigned to this penalty, when the ESG score is too high so that the portfolio finds a balance between being financially efficient and having a good ESG score.

3.5 ESG Ratings Providers

In order to use the ESG factors in an investment strategy, these factors need to be translated into ESG scores. The firms that use the disclosures of the company supplied equity or debt, offering "explicitly or implicitly sustainability metrics and information" in order to determine ESG scores are called ESG Ratings providers.

There are several ESG Ratings providers but they do not provide the same ESG scores because each ESG ratings provider has its own methodology.

The most famous providers are MSCI, Sustainalytics, Bloomberg, Thomson Reuters and RobecoSAM. Only the two former will be explained here because they are the most frequently used and the most accessible ones. One thing to know is that there are two main impacts to consider in terms of ESG. The impact of company's activities on the environment and/or on the society and the impact of ESG on the economic value of a company and also, on its financial risk and return profile from an investment perspective. The last impact is considered as relevant in the ESG Risk Ratings and is a material risk if the decisions of a reasonable investor change because of that risk.

3.5.1 MSCI ESG Ratings

The ESG ratings provided by MSCI are expressed as letters from CCC (the worst rating) to AAA (the best one). In order to get these ESG ratings, the methodology of MSCI consists of taking the weighted average of "individual Environmental and Social Key Issue Scores and the Governance Pillar Score is calculated and then normalized relative to ESG Rating industry peers" (MSCI ESG Research LLC, 2022).

And concretely, how does it work?

MSCI has exposed several Key Issues that are classified according to three pillars: Environmental, Social and Governance Pillars. Once Key Issues are identified by MSCI, they are assigned to each GICS sub-industry according to their relevance. Then, weights of each Environmental and Social Key Issue and the weight of the Governance Pillar are determined for each industry.

Concerning the Environmental and Social Pillars, the weights are determined based on the negative and positive impact the industry has on the environment or the society but also based on the expected time frame for risk/opportunity to materialize.

Concerning the Governance Pillar, the weight takes into account the level of contribution to Corporate Governance and Corporate Behavior and also the expected time horizon of risk/opportunities. However, for the former there is only the choice between high and medium contribution and for the latter, there is no choice, it is the long-term time horizon.

Once the weights for each ESG Pillar are found for each industry, the ESG risk is assigned to each industry. In order to get the ESG ratings for each company, management strategies and exposure to ESG risks of each company are examined. Risk exposure is scored from 0 (no exposure) to 10 (very high exposure). And risk management is scored also on a 0-10 scale where 0 means no evidence of management efforts and 10 very big management efforts. The exposure and management risks are then combined to express the fact that if a company is highly exposed to risk, its management efforts also need to be high and inversely. The overall key issue is then also scored on a 0-10 scale. And they do approximately the same for opportunities.

Therefore, for each company a weighted average is done based on risk exposure and risk management of each ESG pillars. Then, it is necessary to adjust this Weighted Average Key Issue Score relative to industry peers and we obtain the Final Industry Adjusted Score that is finally translate into letter ratings (CCC to AAA) by dividing its 0-10 scale into seven equal parts (MSCI, 2022).

3.5.2 Sustainalytics ESG Ratings

The ESG ratings provided by Sustainalytics is quantitative and can be classified into five categories: negligible (0-10), low (10-20), medium (20-30), high (30-40) and severe (40+). The methodology of Sustainalytics consists of taking the "sum of the individual material ESG issues' unmanaged risk scores" (Sustainalytics, 2021).

The unmanaged risk is composed by two different risks: unmanageable risk which occurs when company initiatives do not address that risk and management gap which is the risk that a company could manage but that Sustainalytics assessed as insufficiently managed.

The first step of this methodology is to compute the company exposure which is the sub-industry exposure times the issue Beta. The issue Beta is the deviation of the company's exposure to a material ESG issue from the sub-industry average exposure.

The second step is to compute the manageable risk which is the multiplication of the company's exposure and the manageable risk factor (MRF). The MRF is "the share of risk that is manageable vs. the share of risk that is unmanageable" (Sustainalytics, 2021).

The third step is then to compute the managed risk which is the manageable risk multiplied by the management score rate. The latter is the percentage that measures a company's management of ESG issues.

And the final step is to compute the unmanaged risk which is the difference between the company's exposure and the managed risk. And therefore, the final ESG risk ratings is the sum of each material ESG issues' unmanaged risk scores (Sustainalytics, 2021).

Chapter 4

Minimum-Divergence Portfolio

Generally, the portfolio-selection problem has three steps. First, choose an objective function that corresponds to the investor's preferences. Second, find the optimal portfolio weights by maximizing this objective function. Third, implement an out-of-sample performance analysis to evaluate if the optimal investment strategy is relevant or not.

Concerning the first step, the approach of Markowitz as explained in chapter 1 is appropriate if the asset returns are Gaussian distributed or if investors only care about the two first moments. However, as mentionned in section 2.5, asset returns are not normally distributed. Then higher moments, others than the two first ones, are not equal to zero and are needed to be taken into account.

That is for this reason that Lassance & Vrins (2022) propose a new framework for the first step of the portfolio-selection problem called *Minimum-Divergence portfolio*.

4.1 Main concept

In order to translate investor's preferences, they use a *target-return distribution* because the aim for many investment strategies is to reach a return portfolio distribution with specific performance metrics. And because information about the type of desired return distributions are available, it is possible to model this desired return distribution. For their objective function, Lassance & Vrins (2022) decide to minimize the distance between the portfolio-return distribution and the investor's target-return distribution. This objective function has a non-negligible advantage: all moments can be considered, no need to make a choice among moments to consider. Another advantage of this new framework is that it significantly improves higher moments if the target-distribution is well chosen.

This new approach, *Minimum-divergence portfolio*, models preferences via a distribution of returns instead of some return statistics.

In this new framework for the first step of the portfolio-selection problem there are also several steps to follow: model investor's preferences thanks to a target-return density f_T, model the portfolio return density f_P, and find the optimal portfolio by having the portfolio return density f_P as close as possible to target-return density f_T. In other words, by taking the minimal distance between f_P and f_T, which can be modelled thanks to a divergence measure $\langle f_P|f_T\rangle$.

The minimum divergence portfolio can be expressed as follows:

$$w^* = \underset{w\in W}{\arg\min}\langle f_P|f_T\rangle \tag{4.1}$$

4.2 KL divergence measure

Mahalanobis (1936) introduces the concept of distance between probability measures. Kullback & Leibler (1951) introduce the KL divergence.

An important thing to know about the KL divergence is that it exists between f_P and f_T only if f_P is absolutely continuous with respect to f_T. In other words, the support set of f_T, Ω_T must include the one of f_P, Ω_P. If this condition is respected, the KL divergence between f_P and f_T is defined as follows:

$$\langle f_P|f_T\rangle := \mathbb{E}[\ln \frac{f_P(P)}{f_T(P)}] = \int_{\Omega_P} f_P(x)\ln\frac{f_P(x)}{f_T(x)}dx \tag{4.2}$$

Another thing to note is that the KL divergence can be considered as a divergence measure only if it satisfies the important criteria of a distance measure: the positive definiteness. A divergence measure is defined as $D(f_P, f_T)$ and the positive definiteness of this measure is defined as follows: $D(f_P, f_T) \geq 0$ for all f_P and f_T. And $D(f_P, f_T) = 0$ if and only if $f_P = f_T$ almost everywhere.

The KL divergence measure will be used here because it is often used for financial problems (Jiang et al., 2018) and also because it takes into account higher moments, which is in line with the aim to capture investor's preferences.

The KL divergence can be decomposed as follows:

$$\langle f_P | f_T \rangle = \mathbb{E}[\ln f_P(P)] - \mathbb{E}[\ln f_T(P)] = -H[P] - \mathbb{E}[\ln f_T(P)] \tag{4.3}$$

where $H[P]$ is the portfolio-return entropy (Shannon, 1948).

That means that KL divergence measures the amount of lost information when approximating the density f_T by f_P.

4.3 Portfolio Density

In order to implement the KL divergence measure, it is essential to estimate the portfolio density.

There are two main approaches to density estimation: the parametric and the non parametric ones. The parametric technique describes a model up to a small number of parameters and then guesses the parameters using the likelihood principle (Botev et al., 2010). However, because we don't have any information about the portfolio returns, it is difficult to use the parametric approach for the portfolio density estimation. Then, the non parametric approach will be used because less rigorous assumptions regarding the distribution of the observed data will be made (Silverman, 1986).

4.3.1 Kernel Density

Kernel density estimation is the most often used nonparametric density estimation method (Botev et al., 2010). The kernel function K will usually be a symmetric density function and must satisfy the following condition:

$$\int_{-\infty}^{+\infty} K(x)dx = 1 \tag{4.4}$$

The kernel estimator with kernel K is defined by

$$\hat{f}_P(x) = \frac{1}{Nh} \sum_{i=1}^{N} K(\frac{x - x_i}{h}) \tag{4.5}$$

where h is the bandwidth or the smoothing parameter.

Therefore, the kernel estimator is the sum of "bumps" placed at the observations. The kernel function will help to determine the shape of the "bumps" and the smoothing parameter will help to determine their width (Silverman, 1986).

In order to implement this kernel estimator, we will use the Gaussian density function for the kernel density function:

$$\hat{f}_P(x) = \frac{1}{Nh}\sum_{i=1}^{N}\phi(\frac{x - x_i}{h}) \tag{4.6}$$

Therefore, it can be interpreted as an equi-weighted sum of N gaussians centered on each observation i with the same width h.

The issue with this non parametric approach is that the bandwith is the same for each gaussian and the choice of this bandwith therefore has a big impact on the density. If the smoothing parameter h is too small, then the density will be not enough smooth. And if the smoothing parameter h is too big, then the density will be over-smooth and then we will lose information.

4.3.2 Model-Based Clustering Density

In order to have a smoother density function, the model-based clustering can be used. There exists different clustering methods such as partitioning techniques, hierarchical techniques and model-based methods. The latter uses probabilistic distributions to create clusters. The model-based clustering aims to find the number of gaussians, their locations (the mean of each gaussian) and their covariances (Fraley & Raftery, 2002).

There are 14 different types of covariance depending on the volume, the shape and the orientation of each cluster (Scrucca et al., 2016). Thanks to the Bayesian Information Criterion (BIC), the optimal number of clusters can be found.

Based on the optimal number of clusters, the mean and the covariance, the Expectations-Maximization (EM) algorithm can be implemented which is divided into two steps.

The first step is the E-Step where the probability that a data point i is under a j gaussian (or cluster) is computed for each i, j such as:

$$\hat{z}_{i,j} = \frac{\hat{\tau}_j f_j(x_i|\hat{O}_j)}{\sum_{j=1}^{G} \hat{\tau}_j f_j(x_i|\hat{O}_j)} \tag{4.7}$$

The second step is the M-Step that will compute the weight of each gaussian such as:

$$\hat{\tau}_j := \frac{n_j}{n} \tag{4.8}$$

where $n_j := \sum_{i=1}^{n} \hat{z}_{i,j}$.

And also in this step, the parameters will be updated as follows:

$$\hat{\mu}_j := \frac{\sum_{i=1}^{n} \hat{z}_{i,j} x_i}{n_j} \tag{4.9}$$

which is the weighted average of points in a given cluster or gaussian.

And in order to find the portfolio density, we need to do the weighted sum of gaussians :

$$\sum_{j=1}^{G} \hat{\tau}_j \phi(\mu_j, \sigma_j^2) \tag{4.10}$$

4.4 Target Density

Lassance & Vrins (2022) propose to match the mean and variance of the target density to those of a reference portfolio. A reference portfolio needs to be on the mean-variance efficient (MVE) frontier because these portfolios outperform the others in terms of mean and variance.

Then the mean and variance of the target-returns should be chosen as follows:

$$(\mu_T, \sigma_T^2) \leftarrow (\mu_0, \sigma_0^2) = (w_0' \mu, w_0' \Sigma w_0) \tag{4.11}$$

where w_0 represents the reference portfolio.

As target density, the Gaussian distribution has been chosen because it is the most famous distribution and also because its higher-moments, the skewness and the kurtosis, are respectively 0 and 3 (or 0 for excess kurtosis). Then, the upper moments are more desirable than those of returns that are often negatively skewed and leptokurtic.

The target density is then as follows:

$$f_T(x; \mu_T, \sigma_T^2) = \frac{1}{\sigma_T^2}\phi(\frac{x - \mu_T}{\sigma_T^2}) \tag{4.12}$$

Chapter 5

Empirical Analysis

5.1 Data

5.1.1 Assets Data Analysis

In this master thesis, the empirical analysis is based on two different datasets: the assets composing the Euro Stoxx 50 and the Euronext 100. Euro Stoxx 50 is an index which includes the 50 largest market capitalizations in the euro zone as ranked by STOXX Ltd. Euronext 100 is a stock market index of the 100 most highly capitalized companies in the euro zone. We decided to firstly focus on Europe because it is the continent that imposes stricter ESG standards compared to the rest of the world (Howitt, 2022). Both datasets are collected during the period from July 2005 to January 2022 and collected on a daily basis provided 4196 observations. We decided to use daily returns because it is the less likely to be affected by estimation risks (Jagannathan & Ma, 2003). The name of the two different datasets will be *EUR50* for assets taken from the Euro Stoxx 50 and *EUR100* for assets taken from the Euronext 100.These two datasets were downloaded from *Yahoo Finance!*.

Dataset	Name	Time period	N	Frequency	Source
Euro Stoxx 50	EUR50	07/2005 - 01/2022	50	Daily	Yahoo Finance
Euronext 100	EUR100	07/2005 - 01/2022	100	Daily	Yahoo Finance

Table 5.1: Description of the datasets used for the empirical analysis

We also did an analysis of normality of the different assets that compose the dataset thanks to the Shapiro-Wilk test. We used the shapiro.test function in RStudio. This test will test two hypotheses, the null hypothesis that returns are normally distributed and the alternative hypothesis that returns are not normally distributed. If the p-value is greater than a certain alpha level (here 5%), then we fail to reject the null hypothesis. On the other hand, if the p-value is lower than the chosen alpha level, then we reject the null hypothesis. In our case, the p-value was always lower than 5%, so we have to reject the null hypothesis. The returns of different assets are not normally distributed.

5.1.2 Portfolios Analysis

We compare three different types of portfolios: the traditional ones with ESG consideration, which we called ESG traditional portfolios; the traditional portfolios without assets exclusion, which are called benchmark portfolios and the minimum divergence portfolios. The table 5.2 will show the different portfolios used in this master thesis.

Type	Portfolio	Abbreviation
Benchmark Portfolios	Minimum Variance Portfolio	MVPort
	Sharpe Ratio Portfolio	SRPort
	Equally-weighted Portfolio	EWPort
ESG Traditional Portfolios	Minimum Variance Portfolio	ESG MVPort
	Sharpe Ratio Portfolio	ESG SRPort
	Equally-weighted Portfolio	ESG EWPort
Minimum Divergence Portfolios	Gaussian Target return and MVPort as reference portfolio	MD-MVPort
	Gaussian Target return and SRPort as reference portfolio	MD-SRPort
	Gaussian Target return and EWPort as reference portfolio	MD-EWPort

Table 5.2: Description and abbreviation of the different portfolios

5.1.3 ESG Data

The empirical analysis is based on one dataset for the ESG score of companies: the ESG ratings of Sustainalytics. This ESG rating agency provides the ESG Risk Score, then the lower the score, the better.

For the ESG implementation, we only applied the exclusion of assets based on their ESG score. We find the ESG Risk Score from Sustainalytics on their website (Sustainalytics, 2022).

5.2 Methodology

Concerning the methodology for this thesis, we will design ESG-compliant portfolios following two different strategies. The first one is the implementation of benchmark portfolios such as equally-weighted portfolios, Sharpe ratio portfolios and minimum-variance portfolios on a restricted universe considering ESG criteria. And the second one is the implementation of minimum-divergence approach on portfolios with a good ESG score by targeting benchmark portfolios on an unrestricted universe.

Before anything else, because we wanted to apply ESG filter, it is necessary to build datasets from scratch. For each asset composing the Euro Stoxx 50 or the Euronext 100, we selected the stock price for the time-period predetermined. Then, we needed to transform the stock price into asset returns in order to compute the different portfolios. All NA values were replaced by the mean of the asset with missing values. The mean was computed on the all considered period of the dataset. There exists other ways to replace missing values, sometimes there are replaced by zero or the variable/asset is simply removed from the dataset. There are different possibilities to replace these missing values but since they are not very numerous, the choice to replace these values will not have much impact on the result. We also decided not to consider assets with a time-period too small.

Because this master thesis has as objective to see the impact of minimum divergence on moments of ESG-compliant portfolios, we need to implement ESG criteria. We decided to implement them following the exclusion of assets with poor ESG Score technique. The first step was to find the ESG score of each asset that was in the *EUR50* or *EUR100* datasets. We decided not to consider assets for which ESG score was missing for Sustainalytics.

In the first instance, for the exclusion technique, concerning the ESG score of Sustainalytics, only assets with ESG risk score below 20 were kept. Then , in order to observe what is the impact of the ESG filter on the performance, we will keep assets with ESG risk score below 15 and also assets with ESG risk score below 25.

5.2.1 Number of assets in datasets

Concerning the *EUR50* dataset, at the beginning we have 50 assets. However, some assets have recently gone public and therefore have very few observations. In order to have a database with a significant number of observations, we removed assets that were recently listed. We thus arrive at a database composed of 43 assets. We will use this database to calculate the benchmark portfolios.

Then, in order to apply the ESG filter, we will have to remove the assets that do not have an ESG risk score. This gives us 41 assets. Based on these 41 assets, we will apply the ESG filter. For the ESG filter excluding ESG risk scores higher than 20, we have 29 assets left. For the filter excluding ESG risk score higher than 25, we have 35 assets left and finally for the filter excluding ESG risk score higher than 15, we have 23 assets.

For the *EUR100* database, we start with 100 assets, we remove the assets that have been listed too recently. We obtain 73 assets. We calculated the benchmark portfolios on these 73 assets.

For each of these assets there was an ESG risk score. So we applied the ESG filter on these assets and for the filter excluding ESG risk scores above 20, we had only 52 assets. For the filter excluding scores above 25, we had 63 and for the filter excluding scores above 15, we had 43 assets.

5.2.2 First Strategy

The first strategy consists of applying ESG exclusion on both *EUR50* and *EUR100* datasets and find the optimal weights using the minimum-variance portfolio, the Sharpe ratio portfolio and the equally-weighted portfolio.

5.2.3 Second Strategy

The second strategy consists of applying ESG exclusion on both *EUR50* and *EUR100* datasets to form ESG-compliant portfolios. Then, we want to find optimal weights using the minimum-divergence approach. We try to put the distribution of ESG-compliant portfolio as close as possible to the Gaussian distribution using the two first moments of reference portfolios that are not ESG-compliant. The goal is to check if we can achieve better moments with the second strategy compared to the first strategy.

For this strategy, we first find the optimal weights of the benchmark portfolios. We computed the mean and volatility. After that, we wanted to construct a portfolio based on the ESG restricted universe using the exclusion assets technique. We wished to find a way to achieve the traditional portfolios with a portfolio on a ESG restricted universe and to see if it is possible to target the two first moments of these benchmark portfolios but while improving the moments of the ESG traditional portfolios.

We decided to target a Gaussian distribution (which is characterised by a skewness and excess kurtosis equal to zero) for which the mean and the volatility equal the benchmark portfolios' mean and volatility. Therefore, we used the ESG-compliant portfolio and we tried to find optimal weights for which the ESG-compliant portfolio distribution is as close as possible to the Gaussian distribution with the first two moments of the benchmark portfolios. In order to find the smallest distance between these two distributions, we used the KL divergence measure.

5.2.4 Performance Analysis

In order to analyse the performance of the different strategies we decided to do the out-of-sample performance thanks to the rolling window method. We took τ as estimation window daily returns to compute the portfolios. We decided to consider τ as 10 years to perform this rolling window technique. We then computed the portfolio returns on the next six months out of sample and the estimation window is advanced by six months. We repeated this until the end of the sample period. We will also do rolling window with τ equal to 3 and 5 years to evaluate the impact of the sample size on the portfolios' performance.

5.3 Results and Further Analysis

The goal of this master thesis is to see if it is possible while targeting the distribution of a certain portfolio on an unrestricted universe to improve the performance measures of an ESG-compliant portfolio. We try to make a compromise between trying to find the same first two moments as the target distribution but also the same higher order moments. By having non-Gaussian returns, we will accept to be on a portfolio that may not be efficient but that has better higher moments than the target portfolios. However, we will check where the ESG-compliant portfolio using the minimum-divergence approach is located in the mean-variance space (on, below or above the efficient frontier of the ESG-compliant portfolio using the traditional approach).

5.3.1 In-sample performance

In this section, the in-sample performance is useful to see if the minimum-divergence portfolios outperform the traditional portfolios on ESG restricted universe. The results of the in-sample performance regarding the different moments of all portfolios are exposed in the table 5.3.

Type	Portfolios	Mean	Volatility	Skewness	Kurtosis	Ann. SR	Adjusted SR	MVaR
EUR50								
Benchmark Portfolios	MVPort	0.052%	0.660%	-0.040	2.385	1.282	1.061	1.009%
	SRPort	0.070%	0.839%	0.505	6.629	1.365	0.820	1.073%
	EWPort	0.047%	0.842%	-0.539	7.168	0.856	0.603	1.340%
ESG Traditional Portfolios	ESG MVPort	0.050%	0.718%	-0.196	4.850	1.107	0.238	1.149%
	ESG SRPort	0.068%	0.831%	-0.275	7.513	1.292	0.810	1.261%
	ESG EWPort	0.043%	1.188%	-0.414	6.954	0.504	0.449	1.880%
Minimum Divergence Portfolios	MD-MVPort	0.044%	0.903%	-0.208	3.609	0.747	0.665	1.427%
	MD-SRPort	0.045%	0.848%	-0.275	4.021	0.813	0.693	1.346%
	MD-EWPort	0.042%	0.890%	-0.193	4.698	0.719	0.629	1.386%
EUR100								
Benchmark Portfolios	MVPort	0.039%	0.606%	-0.892	10.630	1.013	0.693	0.998%
	SRPort	0.081%	1.032%	-0.328	3.517	1.275	0.733	1.602%
	EWPort	0.043%	0.946%	-0.603	7.106	0.676	0.539	1.533%
ESG Traditional Portfolios	ESG MVPort	0.035%	0.651%	-0.967	12.348	0.847	0.495	1.291%
	ESG SRPort	0.080%	0.876%	-0.150	5.641	1.445	1.061	1.828%
	ESG EWPort	0.044%	1.016%	-0.558	6.425	0.636	0.529	1.650%
Minimum Divergence Portfolios	MD-MVPort	0.048%	0.988%	-0.482	6.719	0.740	0.583	1.573%
	MD-SRPort	0.039%	0.989%	-0.549	5.932	0.578	0.499	1.618%
	MD-EWPort	0.054%	1.077%	-0.502	4.914	0.754	0.618	1.759%

Table 5.3: In-sample performance: Performance measures of the different portfolios for the *EUR50* and *EUR100* datasets with ESG risk score below 20.

Based on the table 5.3, one can see that for benchmark portfolios and ESG traditional portfolios, the in-sample volatility is lower with the minimum-variance strategy than with the others. This is totally logical because the minimization of the volatility on the training set is the objective of this strategy. And based on the table 5.3, one can also see that for benchmark portfolios and ESG traditional portfolios, the in-sample Sharpe ratio is better for the Sharpe ratio strategy which is also consistent since the goal of this strategy is to maximize the in-sample Sharpe ratio. We can draw a first conclusion here, the minimum-variance and the Sharpe ratio strategies are better than the equally-weighted strategy in-sample because they are located on the efficient frontier. However, when there are a lot of parameters to estimate, the strategies tend to overfit the data and then, poorly perform on new future data. That is why in general, the out-of-sample performance is smaller than the in-sample one. We will check this fact in our out-of-sample analysis. Let's first focus on the in-sample analysis.

We will compare different portfolios between them: firstly the comparison between the benchmark portfolios and the ESG traditional portfolios, secondly the comparison between the benchmark portfolios and the minimum-divergence portfolios and finally, the comparison between the ESG traditional portfolios and the minimum divergence portfolios (see table 5.3).

1. Comparison of the benchmark portfolios with the ESG traditional ones:

    (a) *EUR50* Dataset:

    Concerning the *EUR50* dataset, we can see that the means of all benchmark portfolios are higher than the ones of ESG traditional portfolios. The volatility is lower (then better) for benchmark portfolios compared to the ones of traditional portfolios on ESG universe. One can suspect that the efficient frontier of each type of portfolios will not be located at the same place and that the efficient frontier of the benchmark portfolios will be higher on the left in the mean-variance space (see figure 1.1 for mean-variance space visualization) than the efficient frontier of the ESG traditional portfolios.

    Regarding the higher moments, the *MVPort* and *SRPort* outperform the *ESG MVPort* and *ESG SRPort* respectively in terms of skewness and kurtosis. While the *EWPort* is outperformed by the *ESG EWPort* in terms of skewness and kurtosis.

For the annualized Sharpe ratio and the adjusted Sharpe ratio, there is no doubt, the benchmark portfolios outperform the ESG traditional ones. We can draw the same conclusion for the modified value-at-risk, the benchmark portfolios have lower MVaR than the ones of ESG traditional portfolios.

These conclusions confirm that the efficient frontier of the benchmark portfolios not considering ESG criteria is better than the one of portfolios considering this kind of criteria as already mentionned by Pedersen et al. (2021) (see figure 5.1).

(b) *EUR100* Dataset:

Concerning the *EUR100* dataset, the means of the ESG traditional portfolios is outperformed by the ones of the benchmark portfolios (except for one case) and regarding the volatility, the benchmark portfolios outperform the ESG traditional portfolios except for the *ESG SRPort* which has a lower volatility than the one of *SRPort*.

For the higher moments, the skewness of the ESG traditional portfolios are lower than the ones of the benchmark portfolios (except for one case). While it is the opposite for the kurtosis, the benchmark portfolios outperform the ESG traditional portfolios except for one case: *ESG EWPort* has lower kurtosis compared to the one of *EWPort*.

The annualized Sharpe ratios of the benchmark portfolios is better compared to the ones of traditional portfolios on ESG restricted universe (except for one case). One can draw the same conclusion for the adjusted Sharpe ratio except for one case: *ESG SRPort* has higher adjusted Sharpe ratio than the one of *SRPort*.

And finally, for the modified value-at-risk, we can draw the same conclusion as the one drawn for *EUR50* dataset: the MVaR for the benchmark portfolios are lower than the ones of ESG traditional portfolios.

Because the two first moments of benchmark portfolios are better than those of traditional portfolios on ESG restricted universe, we will try to target a Gaussian distribution with the mean and standard deviation of the benchmark portfolios.

The idea is then to target the Gaussian distribution with these first two moments in order to improve the performance measures of the traditional portfolios with ESG considerations.

2. Comparison between the benchmark portfolios and the minimum divergence portfolios:

   (a) *EUR50* Dataset:

   Concerning the *EUR50* dataset, the means of the benchmark portfolios are higher than the ones of the minimum divergence portfolios. Regarding the volatility, the benchmark portfolios also outperform the minimum divergence ones.

   For the higher moments, the skewness of the minimum divergence portfolios are more negative than the ones of benchmark portfolios expect for one case: the *MD-EWPort* has better skewness compared to the *EWPort*. The lack of improvement in terms of skewness is due to the fact that the skewness of *MVPort* and of *SRPort* are very close to or greater than zero. However, due to the target chosen, the minimum divergence portfolios will try to achieve the higher moments of the Gaussian distribution which are zero for skewness and excess kurtosis. Then it is impossible to do better in terms of skewness for a value greater than zero if we target a Gaussian distribution. One idea to solve this issue is to use another target distribution that has a positive value for its skewness.

   If we take a look at the kurtosis, the minimum divergence portfolios outperform the benchmark portfolios except for one case: the *MVPort* has better kurtosis than the *MD-MVPort*.

   For the annualized Sharpe ratio, there is no doubt, the benchmark portfolios outperform the minimum-divergence portfolios. It is the same for the adjusted Sharpe ratio except for the comparison between the *EWPort* and *MD-EWPort*, the one of *MD-EWPort* is better. And regarding the modified value-at-risk, the benchmark portfolios outperform the minimum divergence portfolios for all cases.

(b) *EUR100* Dataset:

Concerning the *EUR100* dataset, the means of the minimum divergence portfolios are higher than the ones of benchmark portfolios except for one case: the *SRPort* outperforms the *MD-SRPort* in terms of mean. The volatility of the benchmark portfolios are lower than the ones of minimum divergence portfolios except for one case: the volatility of *MD-SRPort* is better than the one of *SRPort*.

For the higher moments, the skewness of the minimum divergence portfolios are higher than the ones of the benchmark portfolios except for the comparison between *SRPort* and *MD-SRPort* where the skewness is better for *SRPort*. For the kurtosis, there is only the *MD-SRPort* that is outperformed by the *SRPort*.

Regarding the annualized Sharpe ratio, the benchmark portfolios do better than the minimum divergence portfolios except for the comparison between the *EWPort* and *MD-EWPort*, the annualized Sharpe ratio is better for the latter. If we look at the adjusted Sharpe ratio, one can draw the same conclusion as for the annualized Sharpe ratio. And finally, for the modified value-at-risk, there is no doubt, the benchmark portfolios outperform the minimum divergence portfolios.

3. Comparison between the ESG traditional portfolios and the minimum divergence portfolios:

(a) *EUR50* Dataset:

Concerning the *EUR50* dataset, the means of the ESG traditional portfolios are higher than the ones of minimum divergence portfolios. For the volatility, the minimum divergence portfolios have lower volatility for *MD-EWPort* than the ESG traditional portfolios *ESG EWPort* while the volatility of *MD-MVPort* and *MD-SRPort* are higher than the ones of *ESG MVPort* and *ESG SRPort*.

For the higher moments, the skewness of the minimum divergence portfolios outperform the skewness of the ESG traditional portfolios except for the *MD-MVPort* which has a higher skewness than the one of *MVPort*. And in terms of kurtosis, it is clear that the minimum divergence portfolios outperform the ESG traditional portfolios in all cases.

While the ESG traditional portfolios outperform the minimum divergence portfolios in terms of annualized Sharpe ratio (except for the comparison between *ESG EWPort* and *EWPort*), the minimum divergence portfolios are better in terms of adjusted Sharpe ratio than the ESG traditional portfolios (except for the comparison between *ESG SRPort* and *SRPort*).

And finally, regarding the modified value-at-risk, the ones of ESG traditional portfolios are better than the ones of minimum divergence portfolios except for the comparison between *ESG EWPort* and *MD-EWPort*.

(b) *EUR100* Dataset:

Concerning the *EUR100* dataset, the means are higher for the minimum divergence portfolios except for the mean of *ESG SRPort* which is higher than the one of *MD-SRPort*. The volatility of the ESG traditional portfolios are lower than the ones of the minimum divergence portfolios without any exception.

For the higher moments, the skewness is better for the minimum divergence portfolios than for the ESG traditional portfolios except for the portfolios computed thanks to the Sharpe ratio strategy. And for the kurtosis, one can see that the ESG traditional portfolios are outperformed by the minimum divergence portfolios except for one case: the *ESG SRPort* has better kurtosis than the *MD-SRPort*.

For the annualized Sharpe ratio, the ESG traditional portfolios outperform the minimum divergence portfolios. However, once again, one can see that there is an exception for the *MD-EWPort* that has better annualized

SR than the ones of *ESG EWPort*. And regarding the adjusted SR, the minimum divergence portfolios outperform the ESG traditional portfolios except for one case.

Finally for the modified value-at-risk, there is only the *MD-SRPort* that outperforms the *ESG SRPort*. For the other portfolios comparison, the ESG traditional portfolios outperform the minimum divergence portfolios.

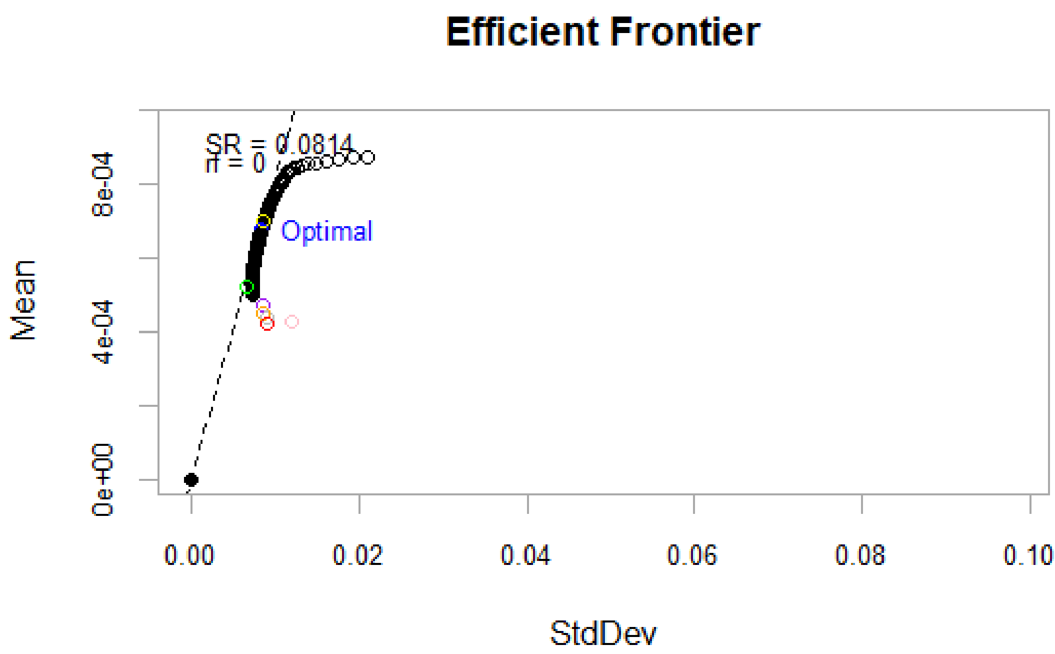
In the table 5.3, one can see, thanks to the mean and volatility of the *MD-SRPort* for each dataset, one specificity of the KL divergence measure. As Lassance & Vrins (2022) observed, the divergence measure will prefer that the variance fit the target distribution rather than the mean. It sanctions a mismatch with the variance of the target distribution more than a mismatch with the mean. And that is for that reason that the mean of the *MD-SRPort* is significantly smaller than the one of the corresponding benchmark portfolio.

On the figure 5.1, one can see that the efficient frontier of the benchmark portfolios is higher than the efficient frontier of the ESG traditional portfolios for the *EUR50* dataset but for the *EUR100* dataset it is not pretty clear. The *MVPort* is better than the *ESG MVPort* on the mean-variance space but the *ESG SRPort* is better than the *SRPort* on this space. This result will impact the performance of the *MD-SRPort* because we are going to target a distribution with two first moments of a portfolio that is outperformed by the ESG-compliant portfolio. Thanks to this figure, we can see that, in terms of mean and variance, the elimination of assets does not have big impact.

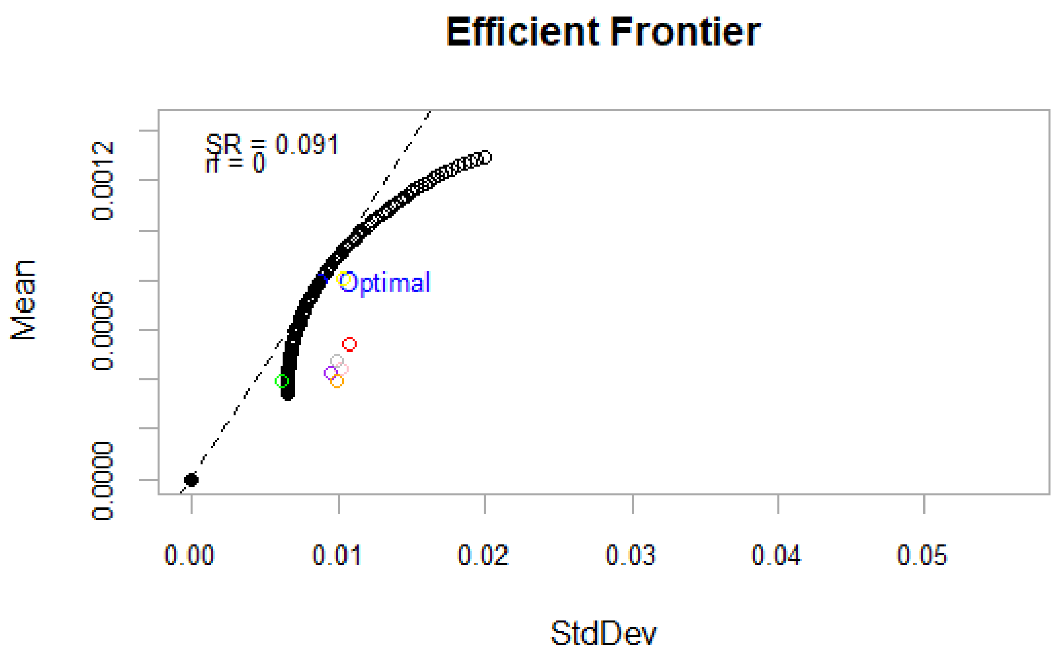
On the mean-variance space, it is logical that the minimum divergence portfolios are worse based on the in-sample results (see the table 5.3).

5.3.2 Out-of-sample performance

In order to confirm the results of the in-sample performance, it could be great to perform an out-of-sample analysis. We follow the rolling window strategy for the two different datasets. We compute the optimal weights on a 10-years period and



(a) Efficient Frontier of the ESG traditional portfolios computed on the *EUR50* dataset.



(b) Efficient Frontier of the ESG traditional portfolios computed on the *EUR100* dataset.

Figure 5.1: Efficient frontiers of the ESG traditional portfolios where additional points added correspond to the other portfolios considered in this thesis. The point in pink corresponds to the *ESG EWPort*, the green one to the *MVPort*, the yellow one to the *SRPort*, the purple one to the *EWPort*, the grey to the *MD-MVPort*, the orange to the *MD-SRPort* and finally, the red one to the *MD-EWPort*.

we keep them constant for the next 6 months. We shift the window, with a length of 10 years, by 6 months, we recalculate the weights and we leave them constant for the 6 months following the estimation window and so on. We assume that 6 months is equal to 126 days. The out-of-sample performance for the different portfolios considered based on the two different datasets will be shown in table 5.4.

Type	Portfolios	Mean	Volatility	Skewness	Kurtosis	Ann. SR	Adjusted SR	MVaR
EUR50								
Benchmark Portfolios	MVPort	0.038%	0.558%	-0.765	6.617	1.091	0.581	0.920%
	SRPort	0.030%	1.219%	-0.599	7.477	0.301	0.283	1.989%
	EWPort	0.043%	0.704%	-0.794	9.969	0.956	0.472	1.124%
ESG Traditional Portfolios	ESG MVPort	0.038%	0.689%	-0.527	6.984	0.855	0.609	1.097%
	ESG SRPort	0.022%	1.237%	0.190	3.663	0.186	0.186	1.853%
	ESG EWPort	0.041%	0.780%	-0.740	9.004	0.817	0.530	1.256%
Minimum Divergence Portfolios	MD-MVPort	0.044%	0.812%	-0.985	12.856	0.834	0.409	1.292%
	MD-SRPort	0.040%	0.850%	-0.673	8.276	0.706	0.529	1.371%
	MD-EWPort	0.038%	0.779%	-0.487	7.461	0.746	0.572	1.230%
EUR100								
Benchmark Portfolios	MVPort	0.034%	0.643%	-1.626	18.656	0.832	0.197	1.047%
	SRPort	0.103%	1.605%	-0.286	3.781	1.002	0.796	2.542%
	EWPort	0.031%	0.950%	-1.504	23.820	0.459	0.310	1.440%
ESG Traditional Portfolios	ESG MVPort	0.018%	0.748%	-1.703	22.061	0.338	0.270	1.199%
	ESG SRPort	0.111%	1.717%	-0.154	4.241	1.009	0.801	2.639%
	ESG EWPort	0.029%	0.997%	-1.486	22.923	0.387	0.295	1.530%
Minimum Divergence Portfolios	MD-MVPort	0.028%	0.955%	-1.951	28.487	0.408	0.273	1.455%
	MD-SRPort	0.032%	1.070%	-1.414	24.832	0.397	0.295	1.582%
	MD-EWPort	0.025%	0.996%	-1.637	24.058	0.326	0.262	1.542%

Table 5.4: Out-of-sample performance: Performance measures of the different portfolios for the *EUR50* and *EUR100* datasets

Before doing any comparison, one can see that for the *EUR50* dataset, the Sharpe ratios are now better for the portfolios constructed thanks to the minimum-variance strategy compared to the ones constructed thanks to the Sharpe ratio strategy. This is due to the fact that the minimum-variance portfolio ignores the mean and because the estimation error of μ is so large (Jagannathan & Ma, 2003), it can lead to better out-of-sample Sharpe ratio. For the benchmark portfolios and ESG traditional portfolios, one can observe that the Sharpe ratios of the portfolios constructed thanks to the minimum-variance strategy are better than the ones of portfolios constructed thanks to the equally-weighted strategy. This is due to the lower out-of-sample volatility for the portfolios constructed thanks to the minimum-variance strategy.

In order to analyse the out-of-sample performance in more details, the different considered portfolios will be compared between them: firstly, the benchmark portfolios with the ESG traditional portfolios, secondly, the benchmark portfolios with the minimum divergence portfolios and finally, the ESG traditional portfolios with the minimum divergence portfolios.

1. Comparison between the benchmark portfolios and the ESG traditional portfolios:

   (a) *EUR50* Dataset:

   Concerning the *EUR50* dataset, the means of the benchmark portfolios outperform the ones of the ESG traditional portfolios because they are higher for the benchmark portfolios. The volatility are also better for the benchmark portfolios because they are lower than the ones of ESG traditional portfolios.

   Regarding the higher moments, the ESG traditional portfolios outperform the benchmark portfolios in terms of skewness and it is the same thing in terms of kurtosis except for one case: the *MVPort* outperforms the *ESG MVPort* in terms of kurtosis.

   For the annualized Sharpe ratio measure, the benchmark portfolios logically outperform the ESG traditional portfolios because its means and volatility were better than the ESG traditional portfolios. However, the adjusted Sharpe ratios of the ESG traditional portfolios outperform the ones of the benchmark portfolios except for the *ESG SRPort* which is lower than the one of the *SRPort*.

   And finally, the MVaR is lower for the benchmark portfolios except for one case: the *ESG SRPort* outperforms the *SRPort* in terms of MVaR.

   (b) *EUR100* Dataset:

   Concerning the *EUR100* dataset, the means of the benchmark portfolios are better than the ones of the ESG traditional portfolios except for one case. For the volatility, the ESG traditional portfolios are outperformed by the benchmark portfolios without any exception.

For the higher moments, the ESG traditional portfolios outperform the benchmark portfolios in terms of skewness except for one case. And regarding the kurtosis, the ESG traditional portfolios are outperformed by the benchmark portfolios except for one case.

The benchmark portfolios are better than the ESG traditional portfolios in terms of annualized Sharpe ratio except for the *ESG SRPort* that has higher annualized SR than the *SRPort*. For the adjusted Sharpe ratio, the type of portfolios that outperforms the other is the ESG traditional type except for one case.

And finally, concerning the modified value-at-risk, the benchmark portfolios outperform the ESG traditional ones.

2. Comparison between the benchmark portfolios and the minimum divergence portfolios:

   (a) *EUR50* Dataset:

   Concerning the *EUR50* dataset, the means of the minimum divergence portfolios are better than the ones of the benchmark portfolios. Only the mean of *EWPort* is higher than the one of *MD-EWPort*. The volatility is generally higher for minimum divergence portfolios except for one case too.

   For the higher moments, only the *MD-EWPort* outperforms the *EWPort* in terms of skewness and kurtosis while the *MVPort* and *SRPort* have better third and fourth moments compared to *MD-MVPort* and *MD-SRPort* respectively.

   The annualized Sharpe ratio measure is also better for benchmark portfolios except for one case. And regarding the adjusted Sharpe ratio, two of three cases are worse for benchmark portfolios compared to the minimum divergence portfolios.

And finally, the MVaR are higher for the minimum divergence portfolios except for *MD-SRPort*'s MVaR which is lower than the one of *SRPort*.

(b) *EUR100* Dataset:

Concerning the *EUR100* dataset, the benchmark portfolios outperform without any exception the minimum divergence portfolios in terms of means. The minimum divergence portfolios are also outperformed in terms of volatility except for the volatility of *MD-SRPort* which is lower than the one of *SRPort*.

For the higher moments, being in terms of skewness or kurtosis, the ones of the benchmark portfolios are better than the ones of minimum divergence portfolios. One can draw the same conclusion for the annualized Sharpe ratios. For the adjusted Sharpe ratio, the benchmark portfolios outperform the minimum divergence ones except for one case and it is the same for the MVaR, the minimum divergence portfolios are outperformed by the benchmark portfolios except for one case.

3. Comparison of the traditional ESG-compliant portfolios and the minimum divergence portfolios:

(a) *EUR50* Dataset:

Concerning the *EUR50* dataset, the means of the minimum divergence portfolios are higher except for one case. The volatility is lower for minimum divergence portfolios except for one case too.

For the higher moments, being for skewness of kurtosis, the ESG traditional portfolios outperform the minimum divergence portfolios except for the *ESG EWPort* which is outperformed by the *MD-EWPort* for both skewness and kurtosis.

Regarding the annualized Sharpe ratio, the ESG traditional portfolios outperform the minimum divergence portfolios except for one case. For the adjusted Sharpe ratio, it is the minimum divergence portfolios that have better results. Finally, for the MVaR, the ESG traditional portfolios are outperformed by the minimum divergence portfolios except for one case: the MVaR of *ESG MVPort* is lower.

(b) *EUR100* Dataset:

Concerning the *EUR100* dataset, the means of the ESG traditional portfolios are better than the ones of minimum divergence portfolios. There is only one exception for the *MD-MVPort* mean which is higher than the *ESG MVPort* one. The volatility is worse for ESG traditional portfolios except for one case.

For the higher moments, there is no doubt, the minimum divergence portfolios are outperformed by the ESG traditional ones for both third and fourth moments. The annualized Sharpe ratios for traditional portfolios on ESG restricted universe are higher than minimum divergence portfolios except for one case. We can conclude the same thing with the adjusted Sharpe ratio and the MVaR measures.

In order to have a clear overview of conclusions about the performances of the portfolio built for strategy 1 and strategy 2, we built the tables 5.5 and 5.6.

Type	Portfolios	Mean	Volatility	Skewness	Kurtosis	Ann. SR	Adjusted SR	MVaR
In-sample performance	***EUR50***							
ESG Traditional Portfolios	ESG MVPort		●	●		●		●
	ESG SRPort	●				●		●
	ESG EWPort	●						
Minimum Divergence Portfolios	MD-MVPort	●			●		●	
	MD-SRPort		●	●	●		●	
	MD-EWPort		●	●	●	●	●	●
Out-of-sample performance	***EUR50***							
ESG Traditional Portfolios	ESG MVPort		●	●	●	●	●	●
	ESG SRPort			●	●			
	ESG EWPort	●				●		
Minimum Divergence Portfolios	MD-MVPort	●						
	MD-SRPort	●	●			●	●	●
	MD-EWPort		●	●	●		●	●

Table 5.5: Summary of best performances between the traditional portfolios on restricted universe and the minimum divergence portfolios with the *EUR50* dataset.

Type	Portfolios	Mean	Volatility	Skewness	Kurtosis	Ann. SR	Adjusted SR	MVaR
In-sample performance	***EUR100***							
ESG Traditional Portfolios	ESG MVPort		●			●		●
	ESG SRPort	●	●	●	●	●	●	
	ESG EWPort		●					●
Minimum Divergence Portfolios	MD-MVPort	●		●	●		●	
	MD-SRPort							●
	MD-EWPort	●		●	●	●	●	
Out-of-sample performance	***EUR100***							
ESG Traditional Portfolios	ESG MVPort		●	●	●			●
	ESG SRPort	●		●	●	●	●	
	ESG EWPort	●		●	●	●	●	●
Minimum Divergence Portfolios	MD-MVPort	●				●	●	
	MD-SRPort	●						●
	MD-EWPort		●					

Table 5.6: Summary of best performances between the traditional portfolios on restricted universe and the minimum divergence portfolios with the *EUR100* dataset.

Thanks to the tables 5.5 and 5.6, we can have a clearer view of the in-sample and out-of-sample of the different portfolios considered in this thesis for both datasets.

Concerning the *EUR50* dataset, one can clearly see that in-sample the minimum divergence portfolios outperform in general the ESG traditional ones in terms of skewness, kurtosis and adjusted Sharpe ratio. However, when we want to check it out-of-sample, there is only the *MD-EWPort* that outperforms the *ESG EWPort* in terms of higher moments. This is strange because the minimum divergence portfolios in theory should improve the higher moments and because in-sample it was pretty clear that the minimum divergence portfolios improve the higher moments compared to the ESG traditional portfolios. In a next section, we will try with other sample sizes for the estimation window in order to see if it has an impact on the results or not.

Concerning the *EUR100* dataset, for the in-sample performance, in most cases, the skewness and kurtosis of the minimum divergence portfolios are better than the ones of ESG traditional portfolios. And we can draw the same conclusion for the asjuted Sharpe ratio. However, for the out-of-sample performance, it is not conclusive for the improvement in higher moments.

The fact that the out-of-sample performance is worse than the in-sample is due to the estimation errors. As Lassance & Vrins (2021) noted, the number of the assets and the moment order determine the number of parameters to estimate. Therefore, when we work on a dataset with many assets, the risk of estimation error is much higher than with a smaller dataset. That can explain the very poor out-of-sample performance for the *EUR100* dataset.

5.3.3 Change in ESG Filter

In this subsection, we are going to analyze the effect of the ESG filter's change on the performance for the different portfolios computed on the two different datasets. Two additional ESG filters were applied on the different datasets: the exclusion of assets having an ESG risk score above 25 (working with 35 assets for *EUR50* database and 63 assets for *EUR100* database) and the exclusion of assets having an ESG risk score above 15 (working with 23 assets for *EUR50* dataset and 43 assets for *EUR100* database). We are going to analyze the in-sample and out-of-sample performances separately.

1. In-sample performance:

   Concerning the tables 5.7 and 5.8, one can see that, whether for the *EUR50* or *EUR100* dataset, the *MD-EWPort* always outperforms the *ESG EWPort* and the *EWPort* in terms of mean, skewness, kurtosis, annualized Sharpe ratio and adjusted Sharpe ratio. And also whether for *EUR50* or *EUR100* dataset, the *ESG SRPort* always outperforms the *MD-SRPort* in terms of mean, skewness, kurtosis, annualized Sharpe ratio and adjusted Sharpe ratio.

   Concerning the higher moments, for both datasets, the *MD-MVPort* and *MD-EWPort* outperform the *ESG MVPort* and *ESG EWPort* in terms of kurtosis and skewness for both ESG filter. And in general, the means of different portfolios decrease when there are more assets taken into account. The ESG traditional portfolios have their means closer to the benchmark portfolios when the assets taken into account increase.

2. Out-of-sample performance:

   The out-of-sample performance (see tables 5.9 and 5.10) of the minimum divergence portfolios is generally poor for both datasets when we use the 10-years estimation window for the rolling window.

   The means of portfolios with the filter that excludes assets with an ESG score above 15 are higher than the ones of portfolios with the exclusion of assets with an ESG score above 25.

   The performance of the ESG traditional portfolios when the number of assets taken into account is bigger is very close to the one of the benchmark portfolios. Then, with an ESG filter that is relaxing, it could be easier for minimum divergence portfolios to outperform the ESG traditional ones because the target-distribution with the two first moments of benchmark portfolios will also express the two first moments of ESG traditional portfolios. Obviously, the choice of target-return distribution should be well done in order to confirm the previous sentence.

   However, when there are many assets it means that we are more exposed to estimation risks. Therefore, we have to find the right balance between relaxing the ESG constraint to get closer to the target which has the first two moments equivalent to the benchmark portfolios but which also roughly correspond to those of the ESG traditional portfolios and not adding too many assets in the analysis to avoid the risk of estimation errors.

5.3.4 Impact of the estimation window's size

In this section, we implement different sizes for the estimation window in order to see if it has an impact on the out-of-sample performance of the portfolios for strategy 1 and strategy 2. We decided to use as size for the estimation window 3 and 5 years in addition to the 10-years estimation window. We did it only on the *EUR50* dataset but we could also do it on the *EUR100*, it does not matter, we will only evaluate if the sample size of the estimation window has an impact on the performance measures. We will compare the out-of-sample performances of the

Type	Portfolios	Mean	Volatility	Skewness	Kurtosis	Ann. SR	Adjusted SR	MVaR
Benchmark Portfolios	MVPort	0.052%	0.660%	-0.040	2.385	1.282	1.061	1.009%
	SRPort	0.070%	0.839%	0.505	6.629	1.365	0.820	1.073%
	EWPort	0.047%	0.842%	-0.539	7.168	0.856	0.603	1.340%
ESG risk score < 15								
ESG Traditional Portfolios	ESG MVPort	0.055%	0.635%	-0.212	3.246	1.375	1.281	1.105%
	ESG SRPort	0.074%	0.730%	-0.128	1.949	1.614	1.742	1.306%
	ESG EWPort	0.049%	0.775%	-0.139	3.348	0.995	0.834	1.203%
Minimum Divergence Portfolios	MD-MVPort	0.054%	0.746%	-0.163	2.176	1.173	0.989	1.174%
	MD-SRPort	0.048%	0.840%	-0.216	2.800	0.897	0.784	1.336%
	MD-EWPort	0.057%	0.844%	-0.125	2.299	1.081	0.936	1.321%
ESG risk score < 25								
ESG Traditional Portfolios	ESG MVPort	0.054%	0.601%	-0.090	4.467	1.425	1.218	1.063%
	ESG SRPort	0.068%	0.818%	-0.146	2.855	1.368	1.018	1.263%
	ESG EWPort	0.047%	0.957%	-0.538	8.053	0.749	0.558	1.512%
Minimum Divergence Portfolios	MD-MVPort	0.047%	0.717%	-0.182	3.379	1.032	0.845	1.120%
	MD-SRPort	0.055%	0.800%	-0.229	3.589	1.101	0.855	1.254%
	MD-EWPort	0.050%	0.810%	-0.221	3.734	0.968	0.793	1.271%

Table 5.7: In-sample performance for ESG filter: Performance measures of the different portfolios for the *EUR50* dataset

different sample size estimation windows for the different portfolios considered (the 3-years and the 5-years estimation window for strategy 1 and strategy 2 portfolios).

The results of these different rolling windows applied on the *EUR50* dataset can be found in the table 5.11.

As observed in the subsections 5.3.1 and 5.3.2, the in-sample higher moments of the minimum divergence portfolios outperform the in-sample higher moments of the ESG traditional portfolios. However, the conclusion is not the same for the out-of-sample analysis. Then, in order to see if the sample size of the estimation window for the rolling window method has an impact, we compare the out-of-sample performances based on the 3-years and 5-years estimation windows.

Concerning the table 5.11, one can see that the Sharpe ratios of the portfolios constructed thanks to the minimum-variance strategy are better than the ones of the portfolios constructed thanks to the Sharpe ratio strategy. This is due to the fact, already explained above, that ignoring the mean can lead to better out-of-

Type	Portfolios	Mean	Volatility	Skewness	Kurtosis	Ann. SR	Adjusted SR	MVaR
Benchmark Portfolios	MVPort	0.039%	0.606%	-0.892	10.630	1.013	0.693	0.998%
	SRPort	0.081%	1.032%	-0.328	3.517	1.275	0.733	1.602%
	EWPort	0.043%	0.946%	-0.603	7.106	0.676	0.539	1.533%
Risk score < 15								
ESG Traditional Portfolios	ESG MVPort	0.040%	0.614%	-0.800	9.607	1.039	0.586	1.226%
	ESG SRPort	0.083%	0.836%	-0.285	2.440	1.585	1.558	1.781%
	ESG EWPort	0.049%	1.013%	-0.462	5.437	0.725	0.598	1.635%
Minimum Divergence Portfolios	MD-MVPort	0.038%	0.857%	-0.650	6.983	0.659	0.529	1.402%
	MD-SRPort	0.057%	1.075%	-0.411	4.193	0.810	0.672	1.743%
	MD-EWPort	0.053%	0.998%	-0.403	4.974	0.811	0.657	1.599%
Risk score < 25								
ESG Traditional Portfolios	ESG MVPort	0.040%	0.599%	-0.956	13.143	1.047	0.387	1.079%
	ESG SRPort	0.081%	0.805%	-0.339	4.575	1.602	1.187	1.676%
	ESG EWPort	0.044%	0.964%	-0.584	6.817	0.684	0.547	1.563%
Minimum Divergence Portfolios	MD-MVPort	0.047%	0.929%	-0.622	6.472	0.775	0.587	1.517%
	MD-SRPort	0.047%	0.997%	-0.561	6.353	0.701	0.564	1.619%
	MD-EWPort	0.054%	1.032%	-0.466	4.722	0.788	0.644	1.678%

Table 5.8: In-sample performance for ESG filter: Performance measures of the different portfolios for the *EUR100* dataset

sample Sharpe ratio because the estimation error of μ is large (Jagannathan & Ma, 2003). And the Sharpe ratios of the portfolios based on minimum-variance strategy are also better than the ones of portfolios based on equally-weighted strategy due to its lower volatility.

1. 3-years estimation window:

   On the table 5.11 for τ equal to 3 years, we can see that the means of the ESG traditional portfolios outperform the ones of the minimum divergence portfolios except for one case. The volatility are higher with the minimum divergence portfolios, only the *MD-SRPort* has a lower volatility than the one of *ESG SRPort*.

   Concerning the higher moments, the *MD-MVPort* and *MD-EWPort* outperform the *ESG MVPort* and *ESG EWPort* in terms of skewness respectively while it is the opposite for *MD-SRPort* and *ESG SRPort*. For the kurtosis, the minimum divergence portfolios outperform the ESG traditional portfolios except

Type	Portfolios	Mean	Volatility	Skewness	Kurtosis	Ann. SR	Adjusted SR	MVaR
Benchmark Portfolios	MVPort	0.038%	0.558%	-0.765	6.617	1.091	0.581	0.920%
	SRPort	0.030%	1.219%	-0.599	7.477	0.301	0.283	1.989%
	EWPort	0.043%	0.704%	-0.794	9.969	0.956	0.472	1.124%
Risk score < 15								
ESG Traditional Portfolios	ESG MVPort	0.043%	0.603%	-0.656	4.904	1.152	0.695	0.996%
	ESG SRPort	0.075%	1.226%	-0.137	3.641	0.955	0.802	1.898%
	ESG EWPort	0.043%	0.709%	-0.492	5.834	0.958	0.669	1.135%
Minimum Divergence Portfolios	MD-MVPort	0.037%	0.786%	-0.381	16.628	0.714	0.429	1.075%
	MD-SRPort	0.030%	1.026%	0.117	12.915	0.398	0.367	1.355%
	MD-EWPort	0.034%	0.797%	-0.363	11.622	0.647	0.490	1.170%
Risk score < 25								
ESG Traditional Portfolios	ESG MVPort	0.039%	0.589%	-0.660	5.512	1.064	0.663	0.970%
	ESG SRPort	0.045%	1.214%	-0.325	7.715	0.509	0.453	1.873%
	ESG EWPort	0.046%	0.717%	-0.666	7.207	1.020	0.586	1.159%
Minimum Divergence Portfolios	MD-MVPort	0.041%	0.738%	-0.751	9.099	0.857	0.526	1.186%
	MD-SRPort	0.044%	0.798%	-0.571	6.789	0.856	0.609	1.283%
	MD-EWPort	0.048%	0.788%	-0.376	8.071	0.948	0.605	1.202%

Table 5.9: Out-of-sample performance for ESG filter: Performance measures of the different portfolios for the *EUR50* dataset

for one case: the kurtosis of *MD-EWPort* is higher than the one of *ESG EWPort*.

The annualized Sharpe ratios are generally better for ESG traditional portfolios while the adjusted Sharpe ratios are generally better for the minimum divergence portfolios.

Regarding the modified value-at-risk, it is the minimum divergence portfolios that is outperformed by the ESG traditional portfolios on two out of three cases.

2. 5-years estimation window:

   Concerning the 5-years estimation window, the means of the minimum divergence portfolios are better than for the traditional portfolios on ESG restricted universe. The most of volatility of the ESG traditional portfolios outperform the ones of the minimum divergence portfolios.

   For the higher moments, the skewness and kurtosis of the traditional portfolios

Type	Portfolios	Mean	Volatility	Skewness	Kurtosis	Ann. SR	Adjusted SR	MVaR
Benchmark Portfolios	MVPort	0.034%	0.643%	-1.626	18.656	0.832	0.197	1.047%
	SRPort	0.103%	1.605%	-0.286	3.781	1.002	0.796	2.542%
	EWPort	0.031%	0.950%	-1.504	23.820	0.459	0.310	1.440%
Risk score < 15								
ESG Traditional Portfolios	ESG MVPort	0.030%	0.680%	-1.617	17.527	0.675	0.328	1.126%
	ESG SRPort	0.118%	1.659%	-0.143	4.237	1.140	0.848	2.535%
	ESG EWPort	0.033%	0.948%	-1.628	22.689	0.491	0.314	1.483%
Minimum Divergence Portfolios	MD-MVPort	0.034%	0.963%	-1.972	32.288	0.501	0.249	1.391%
	MD-SRPort	0.025%	1.074%	-1.461	21.918	0.295	0.250	1.669%
	MD-EWPort	0.038%	1.004%	-1.695	26.443	0.537	0.285	1.507%
Risk score < 25								
ESG Traditional Portfolios	ESG MVPort	0.035%	0.712%	-1.612	19.776	0.742	0.258	1.144%
	ESG SRPort	0.103%	1.605%	-0.286	3.781	1.002	0.796	2.542%
	ESG EWPort	0.031%	0.963%	-1.569	24.859	0.448	0.302	1.455%
Minimum Divergence Portfolios	MD-MVPort	0.028%	0.963%	-1.856	30.278	0.399	0.270	1.413%
	MD-SRPort	0.025%	1.040%	-1.472	24.486	0.313	0.258	1.564%
	MD-EWPort	0.034%	1.000%	-1.624	26.877	0.477	0.294	1.480%

Table 5.10: Out-of-sample performance for ESG filter: Performance measures of the different portfolios for the *EUR100* dataset

considering ESG criteria are better than the ones of the minimum divergence portfolios.

Two out of three cases of the minimum divergence portfolios have better annualized Sharpe ratio and better adjusted Sharpe ratio.

And finally, regarding the MVaR, only the MVaR of the *MD-SRPort* outperforms the one of *ESG SRPort*. For the rest, it is the ESG traditional portfolio type that outperforms the other.

In order to see more easily the results of this subsection, we built the table 5.12.

Based on the table 5.12, the results for the out-of-sample with a lower estimation window for the rolling window method seem better than the ones with a longer estimation window. The out-of-sample analysis for the 3-years estimation window coincides more with the results obtained with the in-sample analysis. Of course, the in-sample outperforms the out-of-sample, that is totally logical.

Type	Portfolios	Mean	Volatility	Skewness	Kurtosis	Ann. SR	Adjusted SR	MVaR
τ = 3 years								
ESG Traditional Portfolios	ESG MVPort	0.058%	0.694%	-0.592	5.912	1.358	0.559	1.112%
	ESG SRPort	0.061%	1.478%	0.111	7.119	0.571	0.522	2.111%
	ESG EWPort	0.061%	0.834%	-0.324	4.469	1.178	0.799	1.310%
Minimum Divergence Portfolios	MD-MVPort	0.055%	0.807%	-0.345	4.514	1.097	0.780	1.276%
	MD-SRPort	0.062%	0.937%	-0.125	3.524	1.046	0.855	1.446%
	MD-EWPort	0.055%	0.865%	-0.306	4.669	1.001	0.755	1.361%
τ = 5 years								
ESG Traditional Portfolios	ESG MVPort	0.052%	0.689%	-0.594	5.910	1.225	0.624	1.110%
	ESG SRPort	0.037%	1.377%	0.066	3.904	0.325	0.320	2.093%
	ESG EWPort	0.051%	0.804%	-0.434	5.624	0.993	0.692	1.276%
Minimum Divergence Portfolios	MD-MVPort	0.055%	0.788%	-0.334	6.475	1.120	0.671	1.211%
	MD-SRPort	0.056%	0.908%	-0.274	5.526	0.969	0.717	1.405%
	MD-EWPort	0.053%	0.833%	-0.497	6.861	0.996	0.631	1.317%

Table 5.11: Out-of-sample performance for different estimation windows: Performance measures of the different portfolios for the *EUR50* dataset

The conclusion of this subsection, then the smaller the sample size the better the out-of-sample, can be explained by the fact that the data are not stationary. In fact, the mean of the returns are not constant over time, then the smaller the sample size of estimation window, the more accurate the estimators.

We did the Augmented Dickey-Fuller test in RStudio in order to see if the data are stationary or not. The hypothesis test confirms that the data are not stationary.

Type	Portfolios	Mean	Volatility	Skewness	Kurtosis	Ann. SR	Adjusted SR	MVaR
τ = 3 years								
ESG Traditional Portfolios	ESG MVPort	•	•			•		•
	ESG SRPort			•				
	ESG EWPort	•	•		•	•	•	•
Minimum Divergence Portfolios	MD-MVPort			•	•		•	
	MD-SRPort	•	•		•	•	•	•
	MD-EWPort			•				
τ = 5 years								
ESG Traditional Portfolios	ESG MVPort		•		•	•		•
	ESG SRPort			•	•			
	ESG EWPort		•	•	•		•	•
Minimum Divergence Portfolios	MD-MVPort	•		•			•	
	MD-SRPort	•	•			•	•	•
	MD-EWPort	•						

Table 5.12: Summary of the out-of-sample performance for different estimation windows: Performance measures of the different portfolios for the *EUR50* dataset

Chapter 6

Conclusion

This master thesis tried to provide analysis about ESG-compliant portfolios using the minimum divergence approach. In this conclusion, we will remind the main results obtained in this thesis, after that we will talk about the limitations of this work and finally suggest ways to continue this research on how to reach a target distribution based on the first moments of reference portfolios in an unconstrained universe with a portfolio considering only assets with a good ESG score.

The first objective was to check if the minimum divergence approach could improve the performance measures of ESG-compliant portfolios by targeting the first two moments of benchmark portfolios. For that objective we performed in-sample and out-of-sample analysis for the benchmark portfolios, ESG traditional portfolios and minimum divergence portfolios. We compared portfolios to each other and we concluded that in-sample, the minimum divergence portfolios generally outperform the ESG traditional portfolios in terms of skewness, kurtosis and adjusted Sharpe ratio for the *EUR50* and *EUR100* datasets. In our different analysis, we noticed that it is difficult to outperform the ESG traditional portfolio constructed thanks to the Sharpe ratio strategy. This is due to the values of skewness and kurtosis of the benchmark portfolio that are already good and then by targeting a Gaussian distribution with skewness and excess kurtosis equal to zero will not help us to improve a slightly positive skewness for example.

Concerning the out-of-sample performance, while we can see only small improvements for the *EUR50* dataset and no improvement for the *EUR100* dataset, we concluded that maybe the sample size of the estimation window could have an impact on the out-of-sample performance. And regarding the *EUR100* database, because there are many assets, we are more exposed to estimation errors and then to poor out-of-sample performance.

In and out-of-sample we also noticed that the mean of *MD-SRPort* is always lower than the ESG traditional portfolios or benchmark portfolios and it is logical because we worked with the KL divergence measure that penalizes more a mismatch with the volatility than with the mean.

Based on our analysis of different thresholds for the ESG filter, we concluded that we need to find the right balance for the number of assets included in the analysis. If the number of assets is closer to the number of assets using to compute the benchmark portfolios, the in-sample performance of the minimum divergence portfolios will be better than the one of ESG traditional portfolios more easily. Once again, the condition of this affirmation is that the target-return distribution must be well chosen. However, if there are more assets taken into consideration, the out-of-sample performance could be poor because of the estimation errors.

And finally, the section where we performed out-of-sample performance on different sample sizes for the estimation window helped us to conclude that the smaller the estimation window size, the better the out-of-sample performance for the minimum divergence portfolios. This can be explained by the non-stationarity of the data. This means that the mean is not constant over time, so the smaller the intervals we take to estimate the parameters, the less estimation errors we will have.

6.1 Limitations

We should note that we were limited by several things that should have impact on our results. Therefore, the results should be analysed with precaution and with consideration of the following aspects:

1. The databases used had to be built by hand, asset by asset. There was no pre-made database that could help us build ESG-compliant portfolios. In fact, if we had used the databases generally used for portfolio analysis, i.e. the databases from K. French's library, it would not have been possible to apply an ESG filter on them because it is not possible to distinguish one company from another. Then it is maybe more prone to errors.

2. The construction of the databases took time. And the runtime for the code

was quite long. This time could not be spent on improving the method, testing different target distributions, testing another way to implement the ESG criteria, making the results more meaningful.

3. Transaction costs were not taken into account even though they cannot be ignored in reality.

4. Due to time constraints, we were only able to test the methods on two different datatsets, which does not represent reality.

5. For the benchmark portfolios, we used traditional portfolios as minimum-variance, Sharpe-ratio and equally-weighted portfolios. However, as mentioned above, we worked with datasets built one asset at a time with data found on yahoo finance which are not simulated data but real data. Therefore we are highly exposed to estimation errors and it is important to have robust estimators because the distribution of asset returns is not stable, asset returns are not normally distributed, and the number of observations contains only 15 years (because of companies that went public only a few years ago). It would have been very interesting to use portfolios constructed with more robust estimators (such as risk parity).

6. Concerning the NA values, they were replaced by the mean of the asset with missing values. However, as said in the section evaluating the impact of the sample size of the estimation window for the rolling window method, the data were not stationary. Then, the mean is not constant over time. All these NA values should be replaced by another value than by the mean.

It is important to keep these limitations in mind when analysing the obtained results.

6.2 Suggestions for further research

Based on the results obtained, it could be interesting to investigate different ways to improve the results or go further in the analysis.

1. In this thesis, we only focus ourselves on datasets for Europe because for the moment it is the Europe the strictest in terms of sustainable regulations and impositions. However, it could be interesting to see if the results are the same for datasets in United States while they do not put as much importance on the sustainability for the moment. We may be able to see whether it is indeed the ESG assets that make it possible or not to reach the target distribution of a portfolio without ESG considerations or whether it is a mass effect. As a result of people buying more ESG assets, then they become more attractive in terms of returns.

2. We only focus on the Gaussian target-distribution in this work but it could maybe improve the results if we use another target-distribution, such as a target distribution that have positive skewness.

3. The ESG filter applied is only the exclusion of assets with poor ESG score. However, it deteriorates the performance of the portfolios because it deletes assets that could be important to reach the desired portfolio and the portfolio is less diversified. It could be interesting to implement the penalization for poor ESG score in the objective function instead of exclusion of poor ESG score.

4. It could be interesting to focus only on different financial crisis and see how the different portfolios behave because many authors precise that the ESG-compliant portfolios outperform the portfolios that do not consider ESG criteria during these financially difficult period.

Bibliography

ASX (2021). *MSCI ESG Ratings methodology.* Online https://www.youtube.com/watch?v=dkMJv14h75o.

Bailey, R., Yeo, J., Jiang, L., & Ferguson, A. (2020). *Where is the demand for ESG coming from?.* Online on Marsh McLennan website https://www.marshmclennan.com/insights/publications/2020/may/esg-as-a-workforce-strategy.html.

Bates, M. (2021). *Where is the demand for ESG coming from?.* Online on ESGClarity website https://esgclarity.com/where-is-the-demand-for-esg-coming-from/.

Bazak, S. & Shapiro, A. (2001). Value-at-risk based risk management: Optimal policies and asset prices. *Review of Financial Studies*, 14(2), 371–405. http://dx.doi.org/10.2139/ssrn.204390.

Black, F. (1993). Estimating expected return. *Financial Analysts Journal*, 49(5), 36–38. https://doi.org/10.2469/faj.v49.n5.36.

Botev, Z., Grotowski, J., & Kroese, D. (2010). Kernel density estimation via diffusion. *The Annals of Statistics*, 38(5), 2916–2957. https://doi.org/10.1214/10-AOS799.

Boudt, K., Peterson, B., & Croux, C. (2008). Estimation and decomposition of downside risk for portfolios with non-normal returns. *Journal of Risk*, 11(2), 79–103. http://dx.doi.org/10.2139/ssrn.1024151.

Brest, P. & Born, K. (2013). *When can impact investing create real impact?.* Online on Standford Social Innovation Review website https://ssir.org/up_for_debate/article/impact_investing#.

Broadstock, D., Chan, K., Cheng, L., & Wang, X. (2021). The role of esg performance during times of financial crisis: Evidence from covid-19 in china. *Finance research letters*, 38(1), 101716. https://doi.org/10.1016/j.frl.2020.101716.

Chopra, V. & Ziemba, W. (1993). The effect of errors in means, variances, and covariances on optimal portfolio choice. *The Journal of Portfolio Management*, 19(2), 6–11. https://doi.org/10.3905/jpm.1993.409440.

Clancy, H. (2021). *How ESG scores can affect the cost of credit.* Online on GreenBiz website https://www.greenbiz.com/article/how-esg-scores-can-affect-cost-credit.

Cont, R. (2001). Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, 1, 223–236. http://dx.doi.org/10.1080/713665670.

DeMiguel, V. & Nogales, F. (2009). Portfolio selection with robust estimation. *Operations Research*, 57(3), 560–577. http://dx.doi.org/10.1287/opre.1080.0566.

Dillon, M. (2020). *Shifting Towards Sustainable Consumption Behavior.* Online on Ksapa website https://ksapa.org/shifting-towards-sustainable-consumption-behavior/.

Dorfleitner, G., Utz, S., & Wimmer, M. (2018). Patience pays off – corporate social responsibility and long-term stock returns. *Journal of Sustainable Finance & Investment*, 8(2), 132–157. https://doi.org/10.1080/20430795.2017.1403272.

European Commission (2022a). *Accord de Paris.* Online https://ec.europa.eu/clima/eu-action/international-action-climate-change/climate-negotiations/paris-agreement_fr.

European Commission (2022b). *Corporate sustainability reporting.* Online https://ec.europa.eu/info/business-economy-euro/company-reporting-and-auditing/company-reporting/corporate-sustainability-reporting_en.

European Commission (2022c). *Delivering the European Green Deal.* Online https://ec.europa.eu/clima/eu-action/european-green-deal/delivering-european-green-deal_en.

European Commission (2022d). *EU taxonomy for sustainable activities.* Online https://ec.europa.eu/info/business-economy-euro/banking-and-finance/sustainable-finance/eu-taxonomy-sustainable-activities_en.

European Commission (2022e). *Overview of sustainable finance.* Online https://ec.europa.eu/info/business-economy-euro/banking-and-finance/sustainable-finance/overview-sustainable-finance_fr.

European Commission (2022f). *Sustainability-related disclosure in the financial services sector.* Online https://ec.europa.eu/info/business-economy-euro/banking-and-finance/sustainable-finance/sustainability-related-disclosure-financial-services-sector_en.

Fraley, C. & Raftery, A. (2002). Model-based clustering, discriminant analysis, and density estimation. *Journal of the American Statistical Association*, 97(458), 611–631. https://doi.org/10.1198/016214502760047131.

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

Howitt, R. (2022). *How European and global sustainability standards for corporate reporting can and will converge.* Online on Reuters website https://www.reutersevents.com/sustainability/how-european-and-global-sustainability-standards-corporate-reporting-can-and-will-converge.

Jagannathan, R. & Ma, T. (2003). Risk reduction in large portfolios: Why imposing the wrong constraints helps. *Journal of Finance*, 58(4), 1651–1684. https://www.jstor.org/stable/3648224.

Jiang, L., Wu, K., & Zhou, G. (2018). Asymmetry in stock comovements: An entropy approach. *Journal of Financial and Quantitative Analysis*, 53(4), 1479–1507. http://dx.doi.org/10.2139/ssrn.2487432.

Johnson, S. (2021). *ESG outperformance narrative 'is flawed', new research shows.* Online on Financial Times website https://www.ft.com/content/be140b1b-2249-4dd9-859c-3f8f12ce6036.

Kullback, S. & Leibler, R. (1951). On information and sufficiency. *The Annals of Mathematical Statistics*, 2(1), 79–86. http://www.jstor.org/stable/2236703.

Lassance, N. (2019). *Information-Theoretic Approaches to Portfolio Selection* [Louvain School of Management Doctoral Thesis, UCLouvain, Mons]. Available at SSRN: http://dx.doi.org/10.2139/ssrn.3500947.

Lassance, N., DeMiguel, V., & Vrins, F. (2021). Optimal portfolio diversification via independent component analysis. *Operations Research*, 70(1), 55–72. http://dx.doi.org/10.2139/ssrn.3285156.

Lassance, N. & Vrins, F. (2021). Portfolio selection with parsimonious higher comoments estimation? *Journal of Banking and Finance*, 126(9), 106–115. https://doi.org/10.1016/j.jbankfin.2021.106115.

Lassance, N. & Vrins, F. (2022). *Portfolio selection: A target-distribution approach.* Available at SSRN: http://dx.doi.org/10.2139/ssrn.3893870.

Ledoit, O. & Wolf, M. (2017). Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets goldilocks. *The Review of Financial Studies*, 30(12), 4349–4388. https://doi.org/10.1093/rfs/hhx052.

Mahalanobis, P. (1936). On divergences and information in statistics and information theory. *Proceedings of the National Institute of Sciences of India*, 2(1), 49–55. https://doi.org/10.1109/TIT.2006.881731.

Maillard, S., Roncalli, T., & Teiletche, J. (2010). The properties of equally weighted risk contribution portfolios. *Journal of Portfolio Management*, 36(4), 60–70. https://doi.org/10.3905/jpm.2010.36.4.060.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1), 77–91. https://doi.org/10.2307/2975974.

Meucci, A. (2009). Managing diversification. *Risk*, 22(5), 74–79. AvailableatSSRN:https://ssrn.com/abstract=1358533.

MSCI (2022). *ESG Ratings & Climate Search Tool.* Online https://www.msci.com/research-and-insights/esg-ratings-corporate-search-tool/issuer/.

MSCI ESG Research LLC (2022). *MSCI ESG Ratings Methodology: Executive Summary.* Online https://www.msci.com/documents/1296102/21901542/ESG-Ratings-Methodology-Exec-Summary.pdf.

Pedersen, L. H., Fitzgibbons, S., & Lukasz 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.

Pézier, J. (2004). *Risk and risk aversion.* In Alexander, C. and Sheedy, E. N., editors, The professional risk managers' handbook. PRMIA Publications, Wilmington DE.

Qian, E. (2022). *Risk parity portfolios: Efficient portfolios through true diversification.* Online https://www.panagora.com/assets/PanAgora-Risk-Parity-Portfolios-Efficient-Portfolios-Through-True-Diversification.pdf.

Scrucca, L., Fop, M., Murphy, T., & Raftery, A. (2016). mclust 5: Clustering, classification and density estimation using gaussian finite mixture models. *R J.*, 8(1), 289–317. https://doi.org/10.32614/RJ-2016-021.

Shannon, C. (1948). A mathematical theory of communication. *Bell Systems Technical Journal*, 27(3), 379–423. https://pure.mpg.de/rest/items/item_2383164/component/file_2383163/content.

Sharpe, W. (1994). The sharpe ratio. *The Journal of Portfolio Management*, 21(1), 49–58. https://doi.org/10.3905/jpm.1994.409501.

Silverman, B. (1986). *Density Estimation for Statistics and Data Analysis.* London: Chapman & Hall. http://dx.doi.org/10.1007/978-1-4899-3324-9.

Sustainalytics (2021). *ESG Risk Ratings - Methodology Abstract: Version 2.1.* Online https://connect.sustainalytics.com/esg-risk-ratings-methodology.

Sustainalytics (2022). *Company ESG Risk Ratings.* Online https://www.sustainalytics.com/esg-rating.

Verheyden, T., Eccles, R., & Feiner, A. (2016). Esg for all? the impact of esg screening on return, risk, and diversification. *Journal of Applied Corporate Finance*, 28(2), 47–55. https://doi.org/10.1111/jacf.12174.

Xiong, J. X. & Idzorek, T. (2011). The impact of skewness and fat tails on the asset allocation decision. *Financial Analysts Journal*, 67(2), 23–35. http://dx.doi.org/10.2469/faj.v67.n2.5.

Zangari, P. (1996). An improved methodology for measuring var. *Risk Metrics Monitor*, 2, 7–25.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
Louvain School of Management
Place des Doyens, 1 bte L2.01.01, 1348 Louvain-la-Neuve
Boulevard Emile Devreux 6, 6000 Charleroi, Belgique
Chaussée de Binche 151, 7000 Mons, Belgique
www.uclouvain.be/lsm